

Convergence of BMI1 and CHD7 on ERK signalling in medulloblastoma.

Sara Badodi¹⁺, Adrian Dubuc²⁺, Xinyu Zhang¹, Gabriel Rosser¹, Mariane Da Cunha Jaeger¹, Michelle M. Kameda-Smith³, Anca Sorana Morrissy², Paul Guilhamon⁴, Philipp Suetterlin⁵, Xiao-Nan Li⁶, Loredana Guglielmi¹, Ashirwad Merve¹, Hamza Farooq², Mathieu Lupien⁴, Sheila K. Singh³, M. Albert Basson⁵, Michael D. Taylor^{2*} and Silvia Marino^{1**}

¹ Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, 4 Newark Street, London E1 2AT, UK

² Program in Developmental & Stem Cell Biology, The Hospital for Sick Children, 101 College Street, TMDT-11-401M, Toronto, ON M5G 1L7, Canada.

³ Pediatric Neurosurgery, Department of Surgery, McMaster Children's Hospital and McMaster Stem Cell & Cancer Research Institute, MDCL 5027, 1280 Main Street West, Hamilton, ON, L8S 4K1, Canada.

⁴ Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada, Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada and Ontario Institute for Cancer Research, Toronto, ON, Canada

⁵ King's College London, Department of Craniofacial Development and Stem Cell Biology, Floor 27, Guy's Hospital Tower Wing, London, SE1 9RT, UK

⁶ Texas Children's Cancer Centre, Texas Children's Hospital, Baylor College of Medicine, 6621 Fannin Street, MC-3-3320, Houston, TX 77479, USA

[^]Corresponding author and Lead contact

Silvia Marino

Blizard Institute

Barts and The London School of Medicine

Queen Mary University of London

4 Newark Street, E1 2AT London, UK

Telephone +442078822585

Email s.marino@qmul.ac.uk

⁺Co-first authors

^{*}Co-senior authors

Summary

We describe a novel molecular convergence between BMI1 and CHD7 in the initiation of medulloblastoma. Identified in a functional genomic screen in mouse models, a BMI1^{High};CHD7^{Low} expression signature within medulloblastoma characterises patients with a poor overall survival. We show that BMI1-mediated repression of the ERK1/2 pathway leads to increased proliferation and tumour burden in primary human MB cells and in a xenograft model, respectively. We provide evidence that DUSP4 repression is dependent on a more accessible chromatin configuration in G4 MB cells with low CHD7 expression. These findings extend the current knowledge of the role of BMI1 and CHD7 in medulloblastoma pathogenesis and raise the possibility that pharmacological targeting of BMI1 or ERK may be particularly indicated in a subgroup of MB with low expression level of CHD7.

Introduction

Medulloblastomas (MB) are the most common malignant brain tumours of childhood. Current standard of care consists of multimodal therapy (surgery with radio- and/or chemotherapy) that does not take into account the specific molecular mechanisms driving tumour growth. Despite a good overall survival rate, long term survivors face very significant treatment-related morbidity and reduced quality of life. Risk stratification of patients based on molecular subtyping and targeted therapies are essential to reduce secondary disabilities and improve overall prognosis.

Gene expression profiling of large cohorts of MB has dissected this historically monomorphous tumour entity into four distinct molecular subgroups (WNT, SHH, Group 3, Group 4) with divergent prognoses and responses to therapy (Taylor et al., 2012). Recently, the heterogeneity within these subgroups has been reduced by two independent studies which have defined additional subtypes (Cavalli et al., 2017; Schwalbe et al., 2017). Whilst WNT, SHH and Group (G)3 MB are well characterised molecularly and have been shown to arise from topographically distinct neural progenitor cells (reviewed in (Leto et al., 2015)), little is known about the molecular and cellular mechanisms underpinning the formation of G4 MB although it is the most frequent (Ramaswamy et al., 2016).

Oncogenic events affecting chromatin-modifying genes have been described across the subgroups (Dubuc et al., 2013), although different modifications are found in the specific subgroups and their frequency varies (Jones et al., 2013). Inactivating mutations in the histone demethylase KDM6A and loss-of-function mutations in the Trithorax gene MLL3 (Northcott et al., 2012) as well as elevated expression of the Polycomb proteins EZH2 (Robinson et al., 2012) and BMI1 (Behesti et al., 2013) were identified particularly in G4 MB,

raising the possibility that increased Polycomb repression is an important step in the pathogenesis of MB and in particular of G4 tumours.

The Polycomb group protein BMI1 is a potent inducer of neural stem cell self-renewal and neural progenitor cell proliferation during development and in adult tissue homeostasis (Leung et al., 2004). Numerous studies demonstrated that BMI1, which is upregulated in a variety of cancers, has a positive correlation with clinical grade/stage and poor prognosis (Glinsky et al., 2005). BMI1 is overexpressed across all MB subgroups with the highest levels of expression detected in G4 tumours, followed by G3, SHH and WNT (Behesti et al., 2013). The growth of G4 MB is dependent on BMI1 expression, as BMI1 knockdown results in reduced tumour growth and invasion in a xenograft model (Merve et al., 2014). Also SHH MB growth is dependent on BMI1 expression, as demonstrated by experimental genetic approaches where crossing of a mouse model of SHH MB onto a *Bmi1*^{-/-} background neutralized tumour formation (Michael et al., 2008). Intriguingly, we recently found that overexpression of Bmi1 alone in the cerebellar granule cell progenitors (GCPs), a cell of origin of MB, is not sufficient to induce MB formation (Behesti et al., 2013). As overexpression of Bmi1 driven by the pan-neural Cre line, Nestin-Cre, did not yield MB formation (Yadirgi et al., 2011), it is conceivable that additional oncogenic events are required to collaborate with BMI1 to induce MB formation.

Here we used a Sleeping Beauty forward genetic approach in the mouse to identify oncogenic events collaborating with Bmi1 over-expression to induce MB formation. We identify a novel molecular convergence between Bmi1 and Chd7, an ATP-dependent chromatin remodeler that fine-tunes developmental gene expression through modulating chromatin structure (reviewed in (Basson and van Ravenswaaij-Arts, 2015)). Importantly, somatic mutations in

CHD7 have been described in G4 MB (Robinson et al., 2012) although their tumourigenicity *in vivo* has not been proven. Here we show that CHD7 controls proliferation of G4 MB cells and its silencing renders them susceptible to Bmi1-mediated ERK1/2-induced proliferation control, raising the possibility that BMI1 and/or ERK may be promising druggable targets in this MB subgroup.

Results

A forward genetic screen identifies Chd7 as a frequent insertion site in MB arising in mice overexpressing Bmi1 in cerebellar progenitors.

The Sleeping Beauty (SB) transposon-based mutagenesis system, which uses a transposase and mutagenic transposon alleles to target mutagenesis to somatic cells of a given tissue in mice was used to identify random mutations capable of driving tumour development. Mice overexpressing Bmi1 in glutamatergic progenitors (*Math1-Bmi1* transgenic line, Fig.S1A,B) were crossed to *SB11; T2Onc2* (Copeland and Jenkins, 2010). Triple mutant mice as well as double mutant controls were generated and kept on tumour watch for 12 months. 10 out of 53 *Math1Bmi1; SB11; T2Onc2* mice developed MB (18.9% incidence) while only 3 MB were detected in 47 *SB11; T2Onc2* mice (6.4% incidence), in keeping with an increased tumorigenicity in the triple mutant mice of 12.5 percentage points (Fig.1A). All tumours had morphological and immunohistochemical characteristics of MB as assessed by Haematoxylin and Eosin (H&E) stain and synaptophysin immunostaining (Fig.1B) and histological features were similar irrespective of their genotype. Next, the tumours were analysed for transposon common insertion sites in genes (gCIS) to identify candidates that could cooperate with Bmi1 in MB development (Supplementary Data S1). A total of 41 recurrent clonal insertions were identified in MB originating in *Math1Bmi1;SB11;T2Onc2*, of which 7 were shared with a published dataset of SB-induced insertion sites in MB occurring in *Ptch1^{+/-}* mice (Morrissy et al., 2016; Wu et al., 2012) (Fig.1C). Among the 41 gCISes identified in *Math1Bmi1;SB11;T2Onc2*, 27 were found to be significantly ($p<0.05$) enriched following Bmi1 overexpression and 3 of these were shared between the two SB-induced MB models (Fig.1D). SB-insertions in *Chd7* (Chromodomain helicase DNA binding factor 7) were inactivating

insertions and so were the majority of insertions in *Ptprd* while the majority of insertions in *Zcchc7* were activating (Fig.1E). *Chd7* is an ATP dependent chromatin remodelling factor (Bajpai et al., 2010), *Ptprd* is a member of the protein tyrosine phosphatase (PTP) family (Pulido et al., 1995) and *Zcchc7* is a component of a nucleolar TRAMP-like complex, an ATP-dependent exosome regulatory complex (Lubas et al., 2011). Notably inactivating insertions in *Chd7* were significantly enriched in the *Math1Bmi1;SB11;T2Onc2* MB model (3/6) as compared to the *Ptch1^{+/-};SB11;T2Onc2* MB model (1/36) ($p < 0.023$) (Fig.1F). Immunostaining for Bmi1 and Chd7 in the SB-induced MB with Chd7 insertions as compared to those without revealed similar and reduced expression respectively (Fig.S1C,D)

Taken together, our data suggests that Bmi1 and Chd7 cooperate to induce MB in a mouse model.

A BMI1^{High};CHD7^{Low} molecular signature identifies a subgroup of G4 MB with reduced overall survival.

Next, we screened primary human medulloblastoma samples for CHD7 mutations. A collection of 830 primary human MB profiled on Affymetrix SNP6.0 copy number arrays and subgrouped using Nanostring technology (Northcott et al., 2012) and 281 primary MB subgrouped using Next-Generation Sequencing (NGS) across three independent cohorts (Jones et al., 2012; Pugh et al., 2012; Robinson et al., 2012) were interrogated. Intriguingly, this analysis identified a molecular convergence on single copy loss of CHD7 via hemizygous deletion (Fig.2A) or hemizygous mutations (Fig.2B) in G4 MB, while no significant

enrichment for CHD7 mutations was found in the SHH and WNT subgroups. A graphical representation of the mutational inactivating events that affect CHD7 is shown in Fig.2C. Next, we examined the subgroup-specific expression levels of BMI1 and CHD7 in a large transcriptomic dataset from primary human MB (n=187, run on Affymetrix HT-HG-U133A arrays) (Cho et al., 2011). While CHD7 is most highly expressed in SHH MBs with relatively reduced expression in G3 and G4 MBs, the opposite patterns are observed for BMI1 (Fig.S2A,B). A Kaplan-Meier survival analysis demonstrates that tumours with BMI1^{High} CHD7^{Low} expression profile have a significantly reduced overall survival as compared to MBs lacking this profile (Fig.2D, p=0.03028). When we examined this relationship in a molecular subgroup-specific manner, a reduced overall survival was observed particularly in G4 MBs with this signature (Fig.2E, p=0.00154). No significant contribution to patients' survival was observed when the relative contribution of BMI1^{High} or CHD7^{Low} was considered (Fig. S2C). These data support the notion that the BMI1/CHD7 molecular convergence identified in mouse models reflects a potentially pathogenically relevant mechanism underpinning a subgroup of G4 MB.

CHD7 controls proliferation of G4 MB cells in vitro in a BMI1-dependent fashion.

Patient-derived MB lines, some of which have been maintained as orthotopic xenografts (Zhao et al., 2012), were chosen to further dissect this novel molecular signature. Firstly, we confirmed the subgroup affiliation by analysing their gene expression against a classifier (Robinson et al., 2012). Principal Component Analysis (PCA) of data obtained from RNAseq analysis confirmed that ICB1299, a line we have recently shown to be dependent on BMI1 for growth and intraparenchymal invasion in a xenograft model (Merve et al., 2014), and

CHLA-01-Med belong to G4 with a significant overlap with G3, whilst ICB1595 belongs to G3 (Fig.3A). These lines were chosen for further analysis and they were expanded and shortly maintained *in vitro* to perform functional studies. Hierarchical clustering of RNA sequencing datasets of the cultured cells and the original datasets obtained from patient's tumour (Zhao et al., 2012) confirmed that although some shift in gene expression did occur after *in vitro* culturing, they remained reasonably similar to the cells isolated from the xenografts and retained their subgroup affiliation (Fig.S3A). Validation on independent cultures confirmed very high expression of Otx2, a transcription factor overexpressed in the majority of G4 MB (Bunt et al., 2010) in all three lines with concomitant high expression levels of the G3/G4 markers (Lin et al., 2016). LMX1A and LHX2 in CHLA-01-Med and EOMEOS and LHX2 in ICB1595 (Fig.3B). The expression of Math1, Gli1 and GAB1, markers of granule cell progenitors and SHH MB (Ayrault et al., 2010) was negligible in these lines compared to two patient-derived SHH MB lines, ICB984 and ICB1338 (Fig.3B).

Next, we compared the expression of BMI1 and CHD7 in these lines with human cerebellum and found similar expression levels for both genes in all MB lines and human cerebellum of early postnatal timepoints (newborn and two months of age). In contrast, the expression of BMI1 was higher in all lines compared to adult cerebellum whilst it was similar for CHD7 (Fig.3C). These findings verified that ICB1299 and CHLA-01-Med are suitable models to test the contribution of CHD7 to G4 MB pathogenesis in the context of BMI1 overexpression. Line ICB1595 was also analysed further to dissect the specificity of the BMI1/CHD7 molecular convergence to G4 MB.

Lentivirus-mediated CHD7 knockdown (Fig.S3B) resulted in increased proliferation of all lines (Fig.3D), an effect that was BMI1-dependent as it was neutralised by concomitant

BMI1 knockdown (Fig.S3B and Fig.3D). Importantly, reconstitution of CHD7 expression in ICb1299 after silencing (Fig. S3D) brought proliferation back to basal level (Fig.S3E,F).

Together, these data indicate that CHD7 represses proliferation in G4 and G3 MB cells that overexpress BMI1.

Genome wide expression analysis identifies CHD7 and BMI1 as key regulators of neuronal differentiation and proliferation of G4 MB primary cells.

To clarify the molecular mechanism mediating the synergistic effect between CHD7 and BMI1 in MB differentiation and proliferation, we carried out a genome-wide analysis of gene expression in the ICb1299 MB line. In particular, we set out to assess the impact of silencing CHD7 in a BMI1 overexpressing G4 MB primary cell line (shCHD7) compared with the same primary cell line treated with scrambled shRNA as control (Ctr vs shCHD7). The impact of silencing BMI1 in shCHD7 cultures compared to silencing only CHD7 was also assessed (shCHD7 vs shCHD7;shBMI1), in addition to an evaluation of the impact of silencing BMI1 in the context of CHD7 expression (Ctr vs shBMI1). 360 genes were differentially expressed in Ctr vs shCHD7 (277 upregulated and 83 downregulated), 1175 in shCHD7 vs shCHD7;shBMI1 (406 upregulated and 769 downregulated) and 584 in Ctr vs shBMI1 (224 upregulated and 360 downregulated) when a threshold of 0.05 for statistical significance and 0.6 for absolute log2 expression change was applied. Firstly, we confirmed that no changes in the subgroup-defining marker expression were noted upon silencing CHD7 or BMI1 or both (Fig.S3C). Pathway analysis carried out on the MetaCore platform highlighted different functions and pathways for these three comparisons. Interestingly, different terms and pathways were impacted in the different conditions tested. In the Ctr vs shCHD7 comparison, Nervous

System Development, Synaptic transmission (GO process), Neuron part and Neuron Projection (GO location) and Synaptic contact, Transmission of nerve impulse, Neurogenesis and Chromatin condensation in prometaphase (process networks) were significantly enriched. Regulation of G1-S transition, Regulation of angiogenesis (process networks), Oxidative phosphorylation and Oxidative stress (Pathway Maps) were highlighted for the comparison shCHD7 vs shCHD7;shBMI1 and cell adhesion (Pathways Maps) and cell cycle S phase and M phase as well as DNA damage were impacted in the Ctr vs shBMI1 comparison (Fig.3E).

Assessment of enrichment for PcG targets among the differentially expressed (DE) genes in CHD7-depleted cells (using the statistical significant cut-off of $p < 0.01$) revealed that 15 of the 68 DE genes were previously reported to possess H3K27me3 on their promoters in mouse neural progenitor cells (Mohn et al., 2008), representing a 2.62-fold enrichment ($p = 0.0035$), (Fig.3F). An enrichment of 2.25-fold ($p = 0.012$) and 4.8-fold ($p\text{-value} < 1e-04$) was found when the data were compared with the proportion of PcG target genes in human fibroblasts (Bracken et al., 2006) and in a MB dataset (Hovestadt et al., 2014) respectively (Fig. 3F). These results support the notion that CHD7 functionally relates to the Polycomb complex in MB.

Taken together, these data suggest that the differentiation potential of MB G4 is controlled by CHD7 via direct or indirect interplays with gene repressed by the Polycomb complex in neural progenitors.

Comparative analysis with G4 MB datasets reveals activation of ERK pathway upon CHD7 silencing via BMI1-mediated repression of DUSP4.

To identify the genes mediating the pro-proliferative effect of CHD7 silencing in a BMI1 overexpressing context and to ensure we were focusing on those relevant for MB patients, a comparative analysis with the expression data of the G4 MB subgroup of patients with a BMI1^{High};CHD7^{Low} signature (Fig.2E) was carried out. Five genes that were upregulated upon BMI1 knockdown in the shCHD7 cell model were also expressed at lower level in the BMI1^{High};CHD7^{Low} patient group compared to the remaining patients, while 15 genes were downregulated in the shCHD7;shBMI1 cell model and expressed at higher levels in the BMI1^{High};CHD7^{Low} patient group (Fig.4A and Fig.S4A). Pathway analysis carried out on the MetaCore platform for all 20 genes indicated ERK signal transduction as the most significant pathway affected (Fig.S4B). A literature review identified DUSP4, HIF-1 α and COL8A2 as being linked to ERK signalling and analysis of a published dataset of ChIPSeq data for Bmi1 in neural stem cells and glioblastoma cells identified DUSP4 as a direct Bmi1 target gene (Gargiulo et al., 2013). Moreover, several members of the DUSP gene family were upregulated in an RNASeq dataset comparing neural stem cells lacking Bmi1 versus control in the mouse (Gargiulo et al., 2013). Upregulation of DUSP4 was confirmed in both G4 MB lines upon silencing of BMI1 (Fig.4B) and lower expression level of DUSP4 was confirmed in the G4 MB subgroup with a BMI1^{High};CHD7^{Low} signature (Fig.4C). Interestingly, no upregulation of DUSP4 expression was noted in the G3 MB line upon CHD7 silencing (Fig.S4C). Next, we assessed chromatin accessibility in the vicinity of the DUSP4 TSS using the Assay for Transposase-Accessible Chromatin (ATAC) method (Buenrostro et al., 2013). Silencing of CHD7 increased DNA accessibility around the DUSP4 TSS and the DUSP4 promoter (Fig.4D,E), a configuration that was lost upon concomitant silencing of BMI1 (Fig.4E).

Western blot analysis for pERK1/2 and ERK1/2 demonstrated pathway activation in conditions where BMI1 is expressed and in a more pronounced fashion when CHD7 is silenced (Fig.5A,B), in keeping with DUSP4 gene repression by BMI1. Indeed, knockdown of BMI1 significantly reduced ERK1/2 pathway activation in G4 MB lines (Fig.5A,B) as predicted, and in line with the upregulation of DUSP4 observed in the genome-wide comparative expression analysis. Importantly, no overactivity of ERK signalling was found in the G3 MB line (Fig.S5A), in keeping with no evidence of DUSP4 upregulation (Fig.S4C). These data suggest a model where CHD7 silencing increases chromatin accessibility at the DUSP4 locus to facilitate BMI1-mediated repression and consequent ERK pathway activation in G4 MB. In conditions where both BMI1 and CHD7 are silenced ERK pathway activation is down to baseline levels, a finding which is unlikely to be mediated by DUSP4 as BMI1/CHD7 are silenced and the chromatin configuration at the locus is closed (Fig.4E).

To determine whether this BMI1-CHD7 connection also applied to freshly isolated cells and not a feature of cells maintained through several *in vitro* passages, we obtained neural progenitor cells (NPC) from human fetal brains. These cells expressed high levels of BMI1 as compared to adult cerebellum, albeit not as high as G4 MB cells (Fig.5C and Fig. S5C). We found that the ERK1/2 pathway activity was increased (Fig.5D,E) upon CHD7 silencing (Fig.S5D), which led to increased proliferation in a BMI1-dependent fashion also in these cells (Fig. 5F).

Next, we assessed whether activation of ERK signalling was an essential mediator of the proliferative phenotype observed upon silencing of CHD7 in G4 MB cells.

Treatment of ICb1299 cells with an ERK inhibitor effectively inhibited the pathway (Fig. S5E) and lead to neutralisation of the increased proliferation observed upon silencing of CHD7,

an effect that was also observed in control cells (Fig.5G and H). Re-expression of CHD7 in ICb1299 rescued the pro-proliferative phenotype induced by CHD7 knockdown (Fig.S3D,E,F) and neutralised the ERK pathway overactivity in these cells (Fig.S5B).

Finally lentiviral-mediated overexpression of mCherry-DUSP4 (Fig. S5F) in ICb1299 abolished the ERK overactivation mediated by CHD7 silencing (Fig. 5 I,L and M).

In conclusion, these data highlight a novel mechanism mediating BMI1/CHD7 pro-proliferative role in MB G4, which impacts on ERK and is regulated by DUSP4.

CHD7 controls tumour volume in a BMI1/pERK-dependent fashion in xenograft models.

To evaluate the *in vivo* relevance of these findings, we injected shCHD7 ICb1299 cells, as well as shCHD7;shBMI1 and shBMI1 cells into the cerebellum of newborn mice. Xenografted mice were kept on tumour watch until symptoms of increased intracranial pressure developed, when they were culled. Histological analysis of the tumours did not reveal morphological differences between MB originating from the different conditions, including the expression of markers such as Synaptophysin (Fig.6A), OTX2 (Fig.S6A) and the lack of expression of GAB1 (Fig.S6B) and GLI1 (Fig.S6C). Stereological assessment of tumour volume revealed larger MBs in mice engrafted with shCHD7 cells, an effect which was lost upon concomitant silencing of BMI1 in the engrafted cells (Fig.6B). Increased proliferation in shCHD7-tumours was confirmed with immunohistochemical detection of Ki67 (Fig.6C,D). Silencing of CHD7 did not affect intraparenchymal invasion, while silencing of BMI1 alone significantly hampered tumour cell invasion (Fig.S6D), in keeping with previous reports (Merve et al., 2014). Overactivity of the ERK pathway was confirmed in the shCHD7 xenografts (Fig.6E).

Together, these data support the conclusion that CHD7 depletion enhances proliferation of G4 MB cells via a BMI1/pERK mechanism.

Discussion

The PcG gene BMI1 is upregulated in a variety of cancers, where it correlates with clinical grade/stage and a poor prognosis. Importantly it is a highly druggable molecule, hence understanding its mechanism of action in tumourigenesis will be essential to guide the further development and clinical testing of BMI1 pharmacological inhibitors.

Here we show that genome wide *in vivo* insertional mutagenesis (T2Onc2) driven by the Sleeping Beauty (SB11) transposase in glutamatergic progenitor cells engineered to over-express Bmi1 results in MB formation. On the Bmi1 over-expressing background, we observe frequent *T2Onc2* inactivating insertions in the chromatin remodelling factor *Chd7* (Bajpai et al., 2010; Whittaker et al., 2017), suggesting that Bmi1 overexpression and Chd7 loss of function cooperate to induce MB. Importantly we show the relevance of this novel molecular convergence for human MB. In fact, high expression of BMI1 in combination with low expression of CHD7 was found to be associated with a poor prognosis in human MB, particularly those of the G4 subgroup and loss of function mutations of *CHD7* are significantly enriched for in this subgroup.

The SB transposon mutagenesis is a powerful *in vivo* screening tool for cancer gene discovery. In the context of MB, this tool has been successfully used to identify novel candidate genes, particularly in the context of SHH MB as expected when using Math1-drivers; however many of the candidate genes identified with this tool are known to be over-expressed or silenced in non-SHH models of the human disease (Dey et al., 2013; Genovesi et

al., 2013; Mumert et al., 2012; Wu et al., 2012). We have studied the BMI1/CHD7 molecular convergence in G4 MB because of the results of the mutation analysis in human tumours indicating that genetic lesions leading to reduced CHD7 expression are found almost exclusively in G4 MB -and never in SHH MB- and because of the negative prognostic correlation of the BMI1^{High};CHD7^{Low} signature found exclusively in G4 MB. Thus, while we initially sought to model the effects and consequences of BMI1-overexpression in a SHH-model of MB, this forward genetic screen uncovered a novel, previously uncharacterized, molecular axis whose signalling modulates the pathogenesis of a proportion of human G4 MB instead. Interestingly, there is mounting evidence of a glutamatergic cell of origin for at least a proportion of G4 MB, as recently suggested by a computational reconstruction of core regulatory circuitry which identified subgroup specific transcription factors (Lin et al., 2016), hence highlighting a potential ontogeny link between the discovery tool we used and the molecular signature we have identified. Further studies in genetically engineered mouse models will be confirm or dispute this hypothesis.

We have used primary G4 MB cells to dissect the contribution of CHD7 to tumourigenesis, both *in vitro* and *in vivo* in a xenograft model. We show deregulated expression of neural differentiation markers upon silencing of CHD7, an effect which is in keeping with a pro-neurogenic role of CHD7 as previously described (Feng et al., 2013; Jones et al., 2015). Genome-wide analysis of gene expression upon CHD7 silencing reveals many more upregulated genes as compared to downregulated genes, in keeping with CHD7 functioning as a repressor in our experimental setting. This is in agreement with a reported role of CHD7 as a repressive modulator of ES-cell specific gene expression via targeting enhancer regions (Schnetz et al., 2010). In keeping with CHD7 functioning as a repressor in

our experimental setup, no synergism between CHD7 and Sox2 was observed, an interaction known to lead to gene activation (Engelen et al., 2011). Sox4 and 11 are also unlikely to mediate CHD7 function in our setting because CHD7 promotes an open chromatin configuration at their promoters (Feng et al., 2013).

We show increased proliferation of G4 MB cells upon CHD7 silencing *in vitro* as well as increased tumour volume upon orthotopic xenografting of these cells into NOD/SCID mice, an effect that is dependent on BMI1 expression. Bioinformatics analysis of the deregulated genes identified in a comparative gene expression analysis with human MB datasets revealed a BMI1-dependent convergence on modulation of ERK signalling via repression of DUSP4 in CHD7 depleted cells. Meta-analysis of publicly available ChIPSeq and RNASeq datasets (Gargiulo et al., 2013) confirmed that DUSP family members are direct target genes of Bmi1. We provide evidence of an altered chromatin accessibility at the promoter of DUSP4 upon CHD7 silencing. This is in keeping with the recently reported ability of CHD7 to alter chromatin accessibility in primary cerebellar granule neuron progenitors (Feng et al., 2017; Whittaker et al., 2017) and it provides support to the notion that PcG-mediated gene repression is counteracted by a compacted chromatin configuration induced by CHD7 in G4 MB cells. Our findings are in good agreement with a previously reported role of DUSP4 in promoting neuronal differentiation and repressing proliferation via repression of the ERK pathway (Kim et al., 2015). Furthermore, activation of ERK signalling has been shown to control a proneural genetic program during cortical development and to activate an Ascl1-controlled neuronal differentiation program (Li et al., 2014). Overactivity of the pathway was found to exert a pro-proliferative role in SHH MB (Brechtel et al., 2014; Sengupta et al., 2012) and in other brain tumours, pilocytic astrocytomas and glioneuronal tumours (Li et al., 2014). It has never been

reported in G4 MB and importantly a biomarker signature potentially predicting a response to ERK inhibitors has never been described. We show that activation of ERK signalling via repression of DUSP4 by BMI1 is not found in a G3 MB line, lending additional support to the conclusion that this molecular convergence is functionally relevant only in a subset of G4 MB. Pharmacological inhibition of ERK signalling confirmed the functional relevance of this novel regulatory loop in the G4 MB lines analysed. Finally, silencing CHD7 in human neural progenitor cells expressing high levels of BMI1, induced overactivation of the ERK pathway and increased proliferation of these cells *in vitro*. These data raise the possibility that a BMI1^{High};CHD7^{Low} signature are oncogenic in human progenitor cells; genetically engineered models recapitulating this signature in a spatio-temporally controlled fashion will be essential to test this hypothesis.

The reversible nature of epigenetic modifications makes them attractive therapeutic targets and pharmacological agents which can potentially reverse altered chromatin states are currently being trialled. Our findings extend the current knowledge of the role of two essential chromatin modifiers, BMI1 and CHD7, in MB pathogenesis and raise the possibility that pharmacological targeting of BMI1 may be particularly indicated in a subgroup of MB with low expression level of CHD7. Alternatively, the use of ERK inhibitors could also be further explored for MB treatment and we provide here essential preclinical evidence that BMI1 and CHD7 could be useful biomarkers to define the patient subgroup that could benefit from this approach.

Experimental procedures

Detailed methods are available in the Supplemental Information.

SB mutagenesis

Male and female Math1-Bmi1, Math1-Bmi1;SB11;T2Onc2 or T2Onc2 mice (12 to 20 weeks of age; at the time they developed signs of MB) were used. We did not perform a formal sample size estimate for the study but based our experimental plan on our previous experience with Sleeping Beauty mutagenesis screening. When mice showed clinical signs of increased intracranial pressure they were culled, the brains were removed and either fixed in formalin for histological assessment or used for genomic DNA extraction. All the procedures involving animals have been approved by the institutional Animal Care Committee, in no case tumour-bearing animals were allowed to reach a tumour burden compromising normal behaviour, food and water intake or exceeding the approved volume of 1,700 mm³.

Culture conditions for patient-derived MB lines

Patient-derived MB lines were obtained from Dr Xiao-Nan Li, Baylor College of Medicine, Texas Children Cancer Centre, USA (Shu et al., 2008; Zhao et al., 2012). ICb1299 were cultured in DMEM (high glucose, GlutaMAX™, ThermoFisher) supplemented with 10% foetal bovine serum and 1% Penicillin-Streptomycin. ICb1595, ICb984 and ICb1338 were cultured in NeuroCult NS-A Basal Medium (Human) (STEMCELL) supplemented with NeuroCult Proliferation Supplement (Human) (STEMCELL), 1% Penicillin-Streptomycin, 2µg/ml Heparin (STEMCELL), 20ng/mL EGF (recombinant mouse, Peprotech) and 10ng/mL bFGF (human recombinant, Peprotech). ICb1595, ICb984 and ICb1338 were cultured on plates coated with Poly-L-ornithine and Laminin (Sigma). CHLA-01-Med were purchased from ATCC (CRL3021) and grown in DMEM-F12 (Gibco) supplemented with B27 (Gibco), 20ng/mL EGF and

10ng/mL bFGF. Cells were maintained at 37°C and were sub-cultured every 3 days once they reached confluence.

Proliferation assay

Patient-derived cells infected with lentiviral constructs coding for shBMI1 and/or shCHD7 were analysed with cell growth curve and EdU staining to assess their proliferation potential compared to scramble control.

***In vivo* orthotopic xenografts**

All procedures had Home Office approval (Animals Scientific Procedures Act 1986, PPL 70/7275) and were carried out as previously reported (Merve et al., 2014). Mice were killed when developing neurological signs and brains were removed and placed in 10% formalin for 24 h and then transferred to PBS.

Inhibition of MEK/ERK activity

Patient-derived cells were treated with the MEK inhibitors, PD98059 (Cell Signaling) or U0126 (Cell Guidance Systems) at different concentrations. After 1 hour of treatment, cells were harvested and lysed for protein analysis. Alternatively, to assess proliferation potential, cells were treated twice with U0126 50 µM before pulsing with EdU or collect them for Western blot analysis.

Western blot analysis

Total protein extracts were prepared with RIPA buffer supplemented with protease inhibitors. Nuclear protein extracts were obtained as previously described (Badodi et al, 2015). Equal amounts of protein extracts were separated by SDS-PAGE and analysed by Western blot.

Statistical Analysis

Differences between groups were determined by one-way ANOVA, followed by Dunnett's post hoc test (ATAC experiments) or Tukey's multiple comparison test (all other experiments) to determine which groups differ significantly from control. Any p-values < 0.05 were considered significant. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

Author contributions

SM and MDT conceived the project; SB, PG, ML, MAB, MDT, SM designed the experiments; SB, XZ, AD, AM, PS, MDCJ, MMKS and MAB conducted the wet lab experiments; SB, XZ, AD, AM, PS, LG, MDCJ, MAB, MDT, MKS, SKS, SM analysed the data; ASM, HF and GR conducted and analysed the computational experiments; SB and SM wrote the paper with contributions from all authors.

Acknowledgements

We are grateful to the BSU staff for help in the daily care of our mouse colony and to the Blizzard Core Facilities (Pathology, Imaging and FACS) for technical advice. This work is supported by grants of the Medical Research Council UK (MR/N000528/1) to SM and MAB, Brain Tumour Research Centre of Excellence to SM, Pathological Society of Great Britain and Ireland (SGS 2013/04/03) to AM, Barts and the London Charity (527/2087) to AM and a CAPES Foundation scholarship, Ministry of Education of Brazil to MDCJ.

References

- Ayrault, O., Zhao, H., Zindy, F., Qu, C., Sherr, C. J., and Roussel, M. F. (2010). Atoh1 inhibits neuronal differentiation and collaborates with Gli1 to generate medulloblastoma-initiating cells. *Cancer research* *70*, 5618-5627.
- Bajpai, R., Chen, D. A., Rada-Iglesias, A., Zhang, J., Xiong, Y., Helms, J., Chang, C. P., Zhao, Y., Swigut, T., and Wysocka, J. (2010). CHD7 cooperates with PBAF to control multipotent neural crest formation. *Nature* *463*, 958-962.
- Basson, M. A., and van Ravenswaaij-Arts, C. (2015). Functional Insights into Chromatin Remodelling from Studies on CHARGE Syndrome. *Trends Genet* *31*, 600-611.
- Behesti, H., Bhagat, H., Dubuc, A. M., Taylor, M. D., and Marino, S. (2013). Bmi1 overexpression in the cerebellar granule cell lineage of mice affects cell proliferation and survival without initiating medulloblastoma formation. *Dis Model Mech* *6*, 49-63.
- Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes & development* *20*, 1123-1136.
- Brechbiel, J., Miller-Moslin, K., and Adjei, A. A. (2014). Crosstalk between hedgehog and other signaling pathways as a basis for combination therapies in cancer. *Cancer Treat Rev* *40*, 750-759.
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., and Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods* *10*, 1213-1218.
- Bunt, J., de Haas, T. G., Hasselt, N. E., Zwijnenburg, D. A., Koster, J., Versteeg, R., and Kool, M. (2010). Regulation of cell cycle genes and induction of senescence by overexpression of OTX2 in medulloblastoma cell lines. *Mol Cancer Res* *8*, 1344-1357.
- Cavalli, F. M. G., Remke, M., Rampasek, L., Peacock, J., Shih, D. J. H., Luu, B., Garzia, L., Torchia, J., Nor, C., Morrissy, A. S., *et al.* (2017). Intertumoral Heterogeneity within Medulloblastoma Subgroups. *Cancer cell* *31*, 737-754 e736.
- Cho, Y. J., Tsherniak, A., Tamayo, P., Santagata, S., Ligon, A., Greulich, H., Berhoukim, R., Amani, V., Goumnerova, L., Eberhart, C. G., *et al.* (2011). Integrative genomic analysis of medulloblastoma identifies a molecular subgroup that drives poor clinical outcome. *J Clin Oncol* *29*, 1424-1430.
- Copeland, N. G., and Jenkins, N. A. (2010). Harnessing transposons for cancer gene discovery. *Nature reviews* *10*, 696-706.
- Dey, J., Dubuc, A. M., Pedro, K. D., Thirstrup, D., Mecham, B., Northcott, P. A., Wu, X., Shih, D., Tapscott, S. J., LeBlanc, M., *et al.* (2013). MyoD is a tumor suppressor gene in medulloblastoma. *Cancer research* *73*, 6828-6837.
- Dubuc, A. M., Remke, M., Korshunov, A., Northcott, P. A., Zhan, S. H., Mendez-Lago, M., Kool, M., Jones, D. T., Unterberger, A., Morrissy, A. S., *et al.* (2013). Aberrant patterns of H3K4 and H3K27 histone lysine methylation occur across subgroups in medulloblastoma. *Acta Neuropathol* *125*, 373-384.
- Engelen, E., Akinci, U., Bryne, J. C., Hou, J., Gontan, C., Moen, M., Szumska, D., Kockx, C., van Ijcken, W., Dekkers, D. H., *et al.* (2011). Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. *Nature genetics* *43*, 607-611.
- Feng, W., Kawauchi, D., Korkel-Qu, H., Deng, H., Serger, E., Sieber, L., Lieberman, J. A., Jimeno-Gonzalez, S., Lambo, S., Hanna, B. S., *et al.* (2017). Chd7 is indispensable for mammalian brain development through activation of a neuronal differentiation programme. *Nat Commun* *8*, 14758.
- Feng, W., Khan, M. A., Bellvis, P., Zhu, Z., Bernhardt, O., Herold-Mende, C., and Liu, H. K. (2013). The chromatin remodeler CHD7 regulates adult neurogenesis via activation of SoxC transcription factors. *Cell stem cell* *13*, 62-72.
- Gargiulo, G., Cesaroni, M., Serresi, M., de Vries, N., Hulsman, D., Bruggeman, S. W., Lancini, C., and van Lohuizen, M. (2013). In vivo RNAi screen for BMI1 targets identifies TGF-beta/BMP-ER stress

pathways as key regulators of neural- and malignant glioma-stem cell homeostasis. *Cancer cell* 23, 660-676.

Genovesi, L. A., Ng, C. G., Davis, M. J., Remke, M., Taylor, M. D., Adams, D. J., Rust, A. G., Ward, J. M., Ban, K. H., Jenkins, N. A., *et al.* (2013). Sleeping Beauty mutagenesis in a mouse medulloblastoma model defines networks that discriminate between human molecular subgroups. *Proceedings of the National Academy of Sciences of the United States of America* 110, E4325-4334.

Glinsky, G. V., Berezovska, O., and Glinskii, A. B. (2005). Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. *J Clin Invest* 115, 1503-1521.

Hovestadt, V., Jones, D. T., Picelli, S., Wang, W., Kool, M., Northcott, P. A., Sultan, M., Stachurski, K., Ryzhova, M., Warnatz, H. J., *et al.* (2014). Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature* 510, 537-541.

Jones, D. T., Jager, N., Kool, M., Zichner, T., Hutter, B., Sultan, M., Cho, Y. J., Pugh, T. J., Hovestadt, V., Stutz, A. M., *et al.* (2012). Dissecting the genomic complexity underlying medulloblastoma. *Nature* 488, 100-105.

Jones, D. T., Northcott, P. A., Kool, M., and Pfister, S. M. (2013). The role of chromatin remodeling in medulloblastoma. *Brain Pathol* 23, 193-199.

Jones, K. M., Saric, N., Russell, J. P., Andoniadou, C. L., Scambler, P. J., and Basson, M. A. (2015). CHD7 maintains neural stem cell quiescence and prevents premature stem cell depletion in the adult hippocampus. *Stem cells (Dayton, Ohio)* 33, 196-210.

Kim, S. Y., Han, Y. M., Oh, M., Kim, W. K., Oh, K. J., Lee, S. C., Bae, K. H., and Han, B. S. (2015). DUSP4 regulates neuronal differentiation and calcium homeostasis by modulating ERK1/2 phosphorylation. *Stem Cells Dev* 24, 686-700.

Leto, K., Arancillo, M., Becker, E. B., Buffo, A., Chiang, C., Ding, B., Dobyns, W. B., Dusart, I., Haldipur, P., Hatten, M. E., *et al.* (2015). Consensus Paper: Cerebellar Development. *Cerebellum*.

Leung, C., Lingbeek, M., Shakhova, O., Liu, J., Tanger, E., Saremaslani, P., Van Lohuizen, M., and Marino, S. (2004). Bmi1 is essential for cerebellar development and is overexpressed in human medulloblastomas. *Nature* 428, 337-341.

Li, S., Mattar, P., Dixit, R., Lawn, S. O., Wilkinson, G., Kinch, C., Eisenstat, D., Kurrasch, D. M., Chan, J. A., and Schuurmans, C. (2014). RAS/ERK signaling controls proneural genetic programs in cortical development and gliomagenesis. *J Neurosci* 34, 2169-2190.

Lin, C. Y., Erkek, S., Tong, Y., Yin, L., Federation, A. J., Zapotka, M., Haldipur, P., Kawauchi, D., Risch, T., Warnatz, H. J., *et al.* (2016). Active medulloblastoma enhancers reveal subgroup-specific cellular origins. *Nature* 530, 57-62.

Lubas, M., Christensen, M. S., Kristiansen, M. S., Domanski, M., Falkenby, L. G., Lykke-Andersen, S., Andersen, J. S., Dziembowski, A., and Jensen, T. H. (2011). Interaction profiling identifies the human nuclear exosome targeting complex. *Mol Cell* 43, 624-637.

Merve, A., Dubuc, A. M., Zhang, X., Remke, M., Baxter, P. A., Li, X. N., Taylor, M. D., and Marino, S. (2014). Polycomb group gene BMI1 controls invasion of medulloblastoma cells and inhibits BMP-regulated cell adhesion. *Acta neuropathologica communications* 2, 10.

Michael, L. E., Westerman, B. A., Ermilov, A. N., Wang, A., Ferris, J., Liu, J., Blom, M., Ellison, D. W., van Lohuizen, M., and Dlugosz, A. A. (2008). Bmi1 is required for Hedgehog pathway-driven medulloblastoma expansion. *Neoplasia (New York, NY)* 10, 1343-1349, 1345p following 1349.

Mohn, F., Weber, M., Rebhan, M., Roloff, T. C., Richter, J., Stadler, M. B., Bibel, M., and Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* 30, 755-766.

Morrissey, A. S., Garzia, L., Shih, D. J., Zuyderduyn, S., Huang, X., Skowron, P., Remke, M., Cavalli, F. M., Ramaswamy, V., Lindsay, P. E., *et al.* (2016). Divergent clonal selection dominates medulloblastoma at recurrence. *Nature* 529, 351-357.

- Mumert, M., Dubuc, A., Wu, X., Northcott, P. A., Chin, S. S., Pedone, C. A., Taylor, M. D., and Fults, D. W. (2012). Functional genomics identifies drivers of medulloblastoma dissemination. *Cancer research* 72, 4944-4953.
- Northcott, P. A., Shih, D. J., Peacock, J., Garzia, L., Morrissy, A. S., Zichner, T., Stutz, A. M., Korshunov, A., Reimand, J., Schumacher, S. E., *et al.* (2012). Subgroup-specific structural variation across 1,000 medulloblastoma genomes. *Nature* 488, 49-56.
- Pugh, T. J., Weeraratne, S. D., Archer, T. C., Pomeranz Krummel, D. A., Auclair, D., Bochicchio, J., Carneiro, M. O., Carter, S. L., Cibulskis, K., Erlich, R. L., *et al.* (2012). Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* 488, 106-110.
- Pulido, R., Krueger, N. X., Serra-Pages, C., Saito, H., and Streuli, M. (1995). Molecular characterization of the human transmembrane protein-tyrosine phosphatase delta. Evidence for tissue-specific expression of alternative human transmembrane protein-tyrosine phosphatase delta isoforms. *The Journal of biological chemistry* 270, 6722-6728.
- Ramaswamy, V., Remke, M., Bouffet, E., Bailey, S., Clifford, S. C., Doz, F., Kool, M., Dufour, C., Vassal, G., Milde, T., *et al.* (2016). Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. *Acta Neuropathol.*
- Robinson, G., Parker, M., Kranenburg, T. A., Lu, C., Chen, X., Ding, L., Phoenix, T. N., Hedlund, E., Wei, L., Zhu, X., *et al.* (2012). Novel mutations target distinct subgroups of medulloblastoma. *Nature* 488, 43-48.
- Schnetz, M. P., Handoko, L., Akhtar-Zaidi, B., Bartels, C. F., Pereira, C. F., Fisher, A. G., Adams, D. J., Flicek, P., Crawford, G. E., Laframboise, T., *et al.* (2010). CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS genetics* 6, e1001023.
- Schwalbe, E. C., Lindsey, J. C., Nakjang, S., Crosier, S., Smith, A. J., Hicks, D., Rafiee, G., Hill, R. M., Iliasova, A., Stone, T., *et al.* (2017). Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. *The Lancet Oncology* 18, 958-971.
- Sengupta, R., Dubuc, A., Ward, S., Yang, L., Northcott, P., Woerner, B. M., Kroll, K., Luo, J., Taylor, M. D., Wechsler-Reya, R. J., and Rubin, J. B. (2012). CXCR4 activation defines a new subgroup of Sonic hedgehog-driven medulloblastoma. *Cancer research* 72, 122-132.
- Shu, Q., Wong, K. K., Su, J. M., Adesina, A. M., Yu, L. T., Tsang, Y. T., Antalffy, B. C., Baxter, P., Perlaky, L., Yang, J., *et al.* (2008). Direct orthotopic transplantation of fresh surgical specimen preserves CD133+ tumor cells in clinically relevant mouse models of medulloblastoma and glioma. *Stem cells (Dayton, Ohio)* 26, 1414-1424.
- Taylor, M. D., Northcott, P. A., Korshunov, A., Remke, M., Cho, Y. J., Clifford, S. C., Eberhart, C. G., Parsons, D. W., Rutkowski, S., Gajjar, A., *et al.* (2012). Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* 123, 465-472.
- Whittaker, D. E., Riegman, K. L., Kasah, S., Mohan, C., Yu, T., Sala, B. P., Hebaishi, H., Caruso, A., Marques, A. C., Michetti, C., *et al.* (2017). The chromatin remodeling factor CHD7 controls cerebellar development by regulating reelin expression. *J Clin Invest.*
- Wu, X., Northcott, P. A., Dubuc, A., Dupuy, A. J., Shih, D. J., Witt, H., Croul, S., Bouffet, E., Fults, D. W., Eberhart, C. G., *et al.* (2012). Clonal selection drives genetic divergence of metastatic medulloblastoma. *Nature* 482, 529-533.
- Yadirgi, G., Leinster, V., Acquati, S., Bhagat, H., Shakhova, O., and Marino, S. (2011). Conditional activation of Bmi1 expression regulates self-renewal, apoptosis, and differentiation of neural stem/progenitor cells in vitro and in vivo. *Stem cells (Dayton, Ohio)* 29, 700-712.
- Zhao, X., Liu, Z., Yu, L., Zhang, Y., Baxter, P., Voicu, H., Gurusiddappa, S., Luan, J., Su, J. M., Leung, H. C., and Li, X. N. (2012). Global gene expression profiling confirms the molecular fidelity of primary tumor-based orthotopic xenograft mouse models of medulloblastoma. *Neuro-oncology* 14, 574-583.

Figure Legends

Figure 1 Identification of oncogenic events cooperating with Bmi1 to promote MB development using the Sleeping Beauty (SB) Transposon System.

A: Kaplan-Meier survival analysis and summary of animal models used to assess the contributions of Bmi1 to MB development. The addition of Bmi1 overexpression to SB11; T2Onc2 led to increased MB formation. **B:** Histology shows morphological (H&E) and immunohistochemical (Synaptophysin) features of MB in Math1-Bmi1; SB11; T2Onc2 **C:** Comparison of recurrent SB-insertions identified in Math1-Bmi1; SB11; T2Onc2 with those identified in Ptch1^{+/-}; SB11; T2Onc2 mice (Morrissey et al., 2016), highlighted CHD7 as a recurrent clonal insertion. **D:** Bar graph representation of recurrent SB-insertions in Math1-Bmi1; SB11; T2Onc2 versus Ptch1^{+/-}; SB11; T2Onc2 mice highlighting events significantly ($p < 0.05$) enriched following Bmi1 overexpression. **E:** SB-insertion maps for candidate genes identified in Math1-Bmi1; SB11; T2Onc2 mice including: *Zcchc7* (activating and inactivating insertions), *Chd7* (inactivating insertions) and *Ptprd* (inactivating and activating insertions). **F:** The statistically significant enrichment of Chd7 insertions identified in Math1-Bmi1; SB11; T2Onc2 MB model versus the Ptch1^{+/-}; SB11; T2Onc2 MB model. Scale bar is 2mm in B left and 125 μ m B centre and right. See also Figure S1.

Figure 2 A BMI1^{High};CHD7^{Low} signature in MB with poor prognosis.

A: Graphical representation of the percentage of MB with CHD7 single copy loss across subgroups with enrichment in G4. **B:** CHD7 mutational analysis across molecular subgroups. **C:** Graphical representation of the mutational inactivating events that affect CHD7 **D-E:** Kaplan-Meier survival analysis demonstrates that tumours with a BMI1^{High};CHD7^{Low} signature have a significantly ($p = 0.03028$) reduced overall survival (D). When we examine this

relationship in a subgroup-specific manner, the same trend ($p=0.00154$) holds true for G4 MB (E). See also Figure S2.

Figure 3 Increased proliferation upon CHD7 silencing is BMI1-dependent in non-WNT, non-SHH MB patient-derived cell models.

A: Scatterplots of the first 3 principal components of the Robinson dataset (Robinson et al., 2012). Points are shaded based on their subgroup. Ellipses represent the 99% confidence intervals of each subgroup, computed using the covariance matrix of each subset. RNAseq data of G4 (ICb1299 and CHLA-01-Med) and G3 primary MB cells were subjected to the same PCA transformation and are shown in cyan, magenta and black respectively. **B:** Heat map representation of relative expression of selected MB markers in patient-derived cells (ICb1299 and CHLA-01-Med: G4; ICb1595: G3; ICb1338 and ICb984: SHH). OTX2, EOMEOS, LMX1A and LHX2 are mainly expressed by G4 and G3 cells while GAB1, GLI1 and MATH1 are specifically expressed by SHH cells. **C:** Expression of BMI1 and CHD7 in ICb1299, CHLA-01-Med and ICb1595 patient-derived MB lines, is comparable to their expression in the early postnatal cerebellum (BMI1) and adult cerebellum (CHD7). **D:** Increased proliferation upon CHD7 silencing is BMI1-dependent ($n=3$). **E:** Pathway analysis reveals different processes and pathways affected by silencing CHD7 and BMI1 individually or concomitantly. **F:** Venn diagram analysis representing the fold enrichment of PcG target genes in the cohort of genes deregulated in CHD7-depleted cells, as compared to the expected values in murine NPC (neural progenitor cells) (Mohn et al., 2008), in human fibroblasts (Bracken et al., 2006) and in human MB (Hovestadt et al., 2014) as previously reported. Data are represented as mean \pm SEM. See also Figure S3.

Figure 4 Molecular convergence on activation of ERK signalling pathway upon CHD7

silencing is mediated by repression of DUSP4 by BMI1.

A: Comparative analysis with human MB datasets identified conserved DE genes in G4 MB cells upon BMI1 knockdown in a CHD7 silenced context and in G4 MB with a BMI^{High};CHD7^{Low} signature. Venn diagram representing the analysis. **B:** Western Blot analysis (top) and quantification (bottom) show CHD7-dependency of BMI1-mediated repression of DUSP4 in G4 MB cells. Histograms show mean \pm SEM (n = 3).

C: Box plot showing lower expression of DUSP4 in BMI1^{High};CHD7^{Low} G4 MB. **D:** Assessment of DNA accessibility in the vicinity of the DUSP4 TSS by Assay for Transposase-Accessible Chromatin (ATAC). Schematic representation of the promoter region of DUSP4 with primer locations. **E:** Increased accessibility in CHD7-silenced cells is lost upon concomitant silencing of BMI1. Data are represented as mean \pm SEM. See also Figure S5.

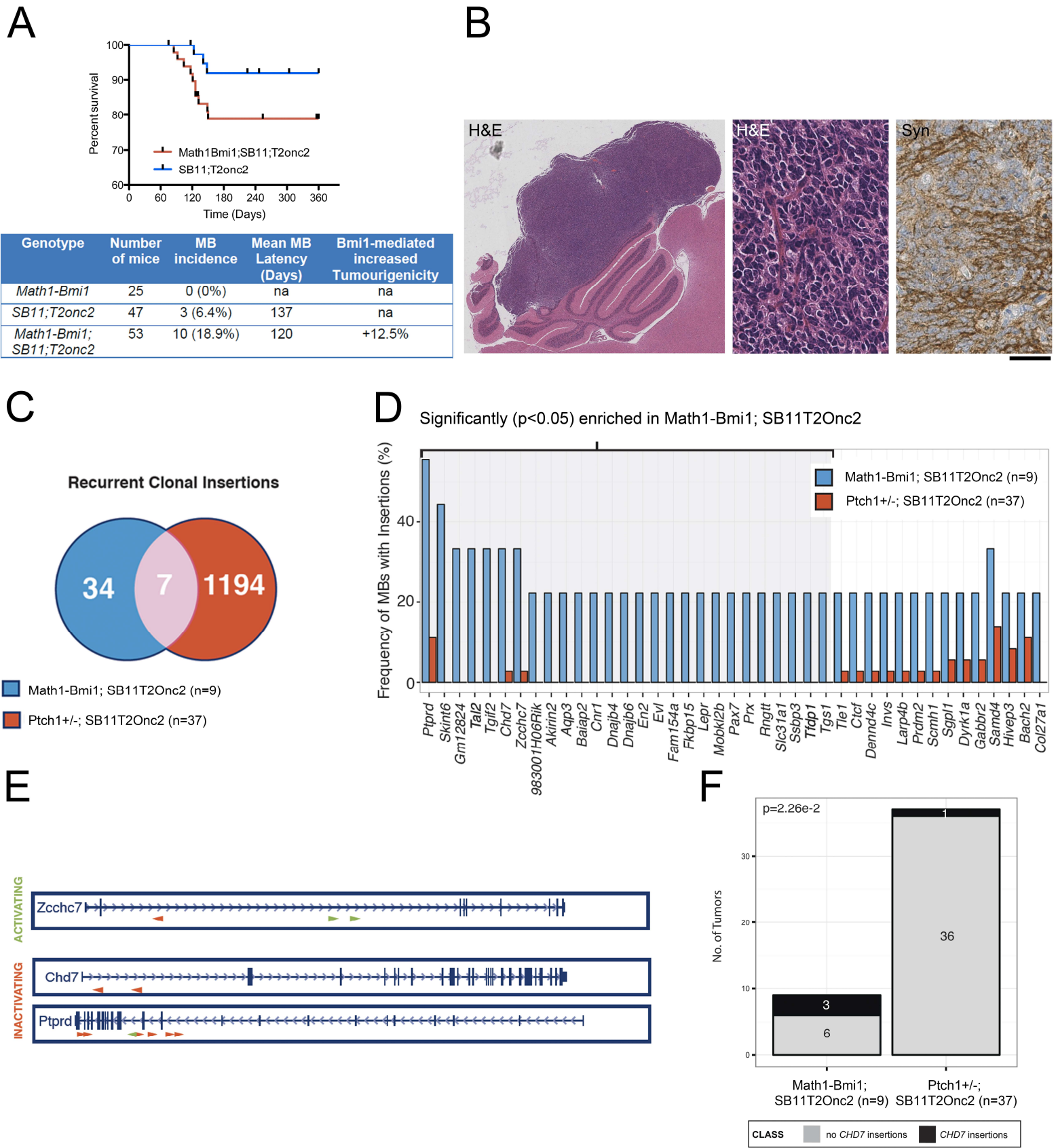
Figure 5 The pro-proliferative state of the G4 MB cells upon silencing of CHD7 is mediated by ERK1/2.

A: Overactivity of the ERK1/2 signalling pathway in ICB1299 CHD7-silenced cells is neutralised by concomitant BMI1 silencing. **B:** Similar effects are seen in CHLA-01-Med. **C:** BMI1 expression in G4 MB cells and in NSC197 is higher than adult cerebellum as shown by Western Blot analysis. **D-E:** Silencing of CHD7 in NPC cells leads to ERK1/2 pathway overactivation as shown by Western blot analysis and quantification (n = 2). **F:** Increased proliferation is seen upon CHD7 silencing in NSC, a finding that is neutralised by incubation with a BMI1 inhibitor **G:** Increased proliferation as assessed by EdU pulse/chase is dependent on ERK1/2 overactivity, as demonstrated by its reduction upon pharmacological inhibition of ERK in ICB1299. **H:** Quantification of the findings. **I-L:** Overexpression of

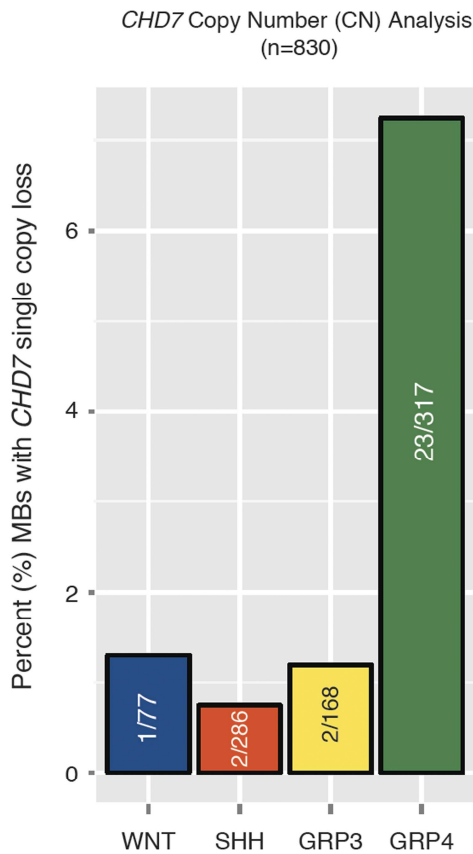
exogenous mCherry-DUSP4 abolishes ERK signalling overactivation mediated by CHD7 silencing. Western blot analysis (H) of ICb1299 transduced with empty LeGO control or LeGO-mCherry-DUSP4 and quantification (I). Data are represented as mean \pm SEM (n=3). **M:** Immunofluorescence of ICb1299 transduced as described in H. Green arrows indicates CHD7 silenced cells with ERK overactivation, while red arrow indicates a cell overexpressing DUSP4 and devoid of ERK phosphorylation. Scale bar is 25 μ m. See also Figure S6.

Figure 6 Increased tumour volume in CHD7-silenced xenografts of G4 MB is BMI1-dependent.

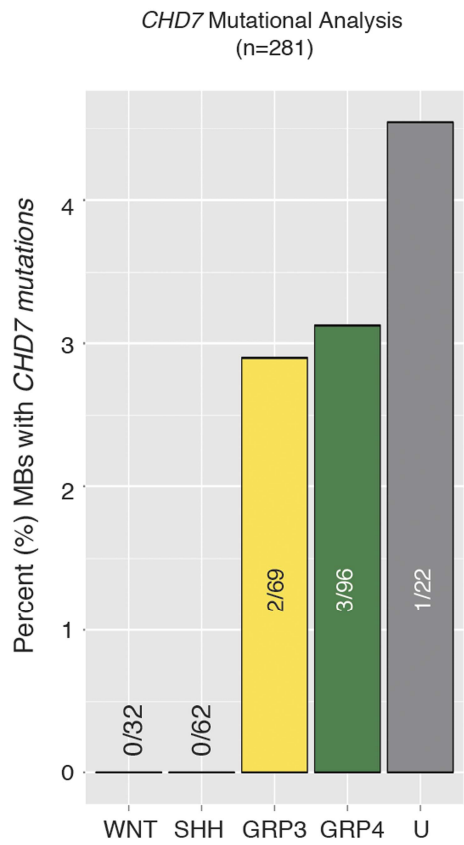
A: Orthotopic xenografts of G4 MB cells silenced for CHD7 or CHD7;BMI1 or BMI1 show similar morphology and they express Synaptophysin. **B:** Increased tumour volume upon injection of shCHD7 G4 MB cells, which is neutralised by concomitant silencing of BMI1. **C** Increased proliferation, as assessed by Ki67 staining, in shCHD7 xenografts is shown together with its BMI1-dependency. **D:** Quantification of the findings. **E:** Overactivation of the ERK pathway is seen in shCHD7 xenografts. Scale bar is 2mm in A and 125 μ m in C and E.



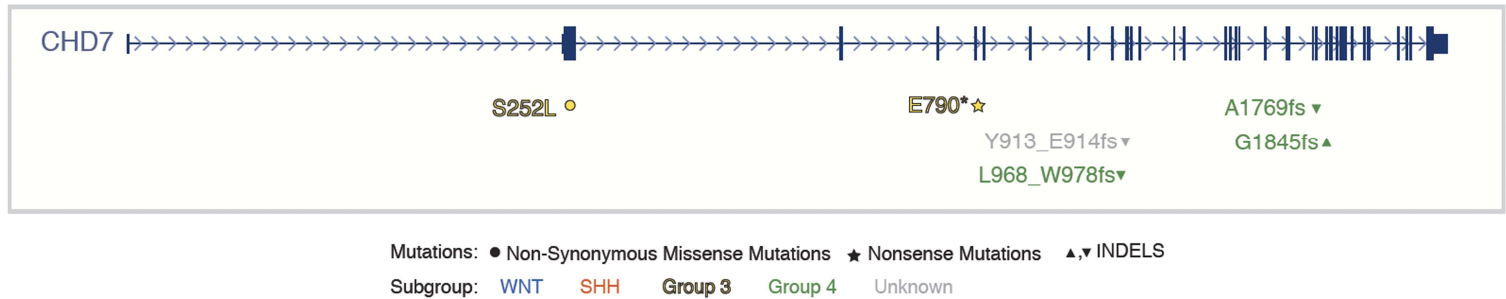
A



B

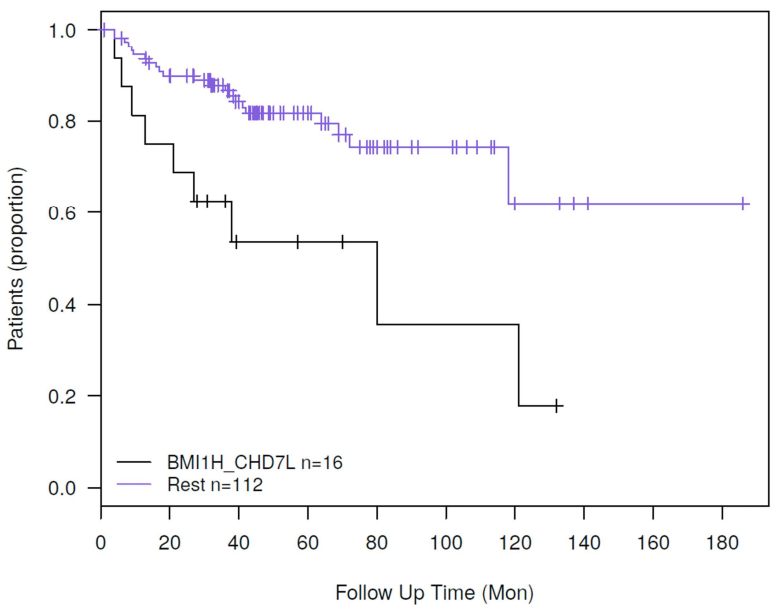


C



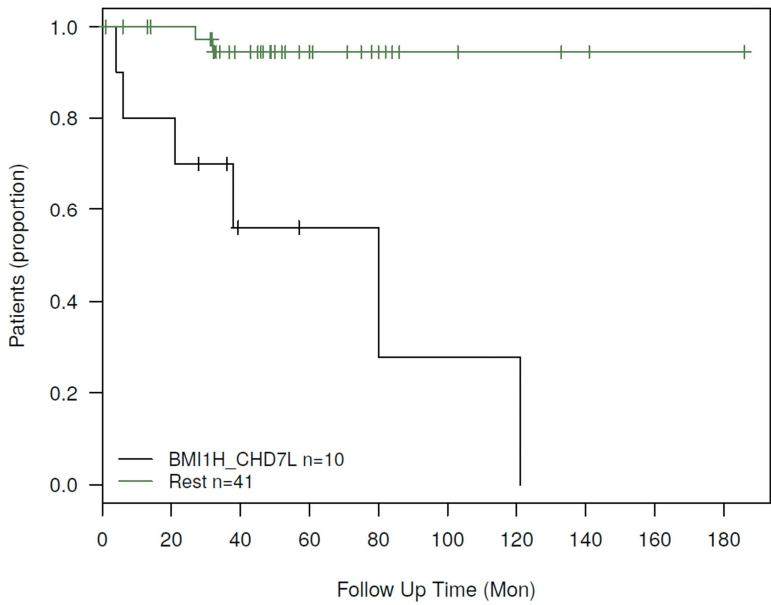
D

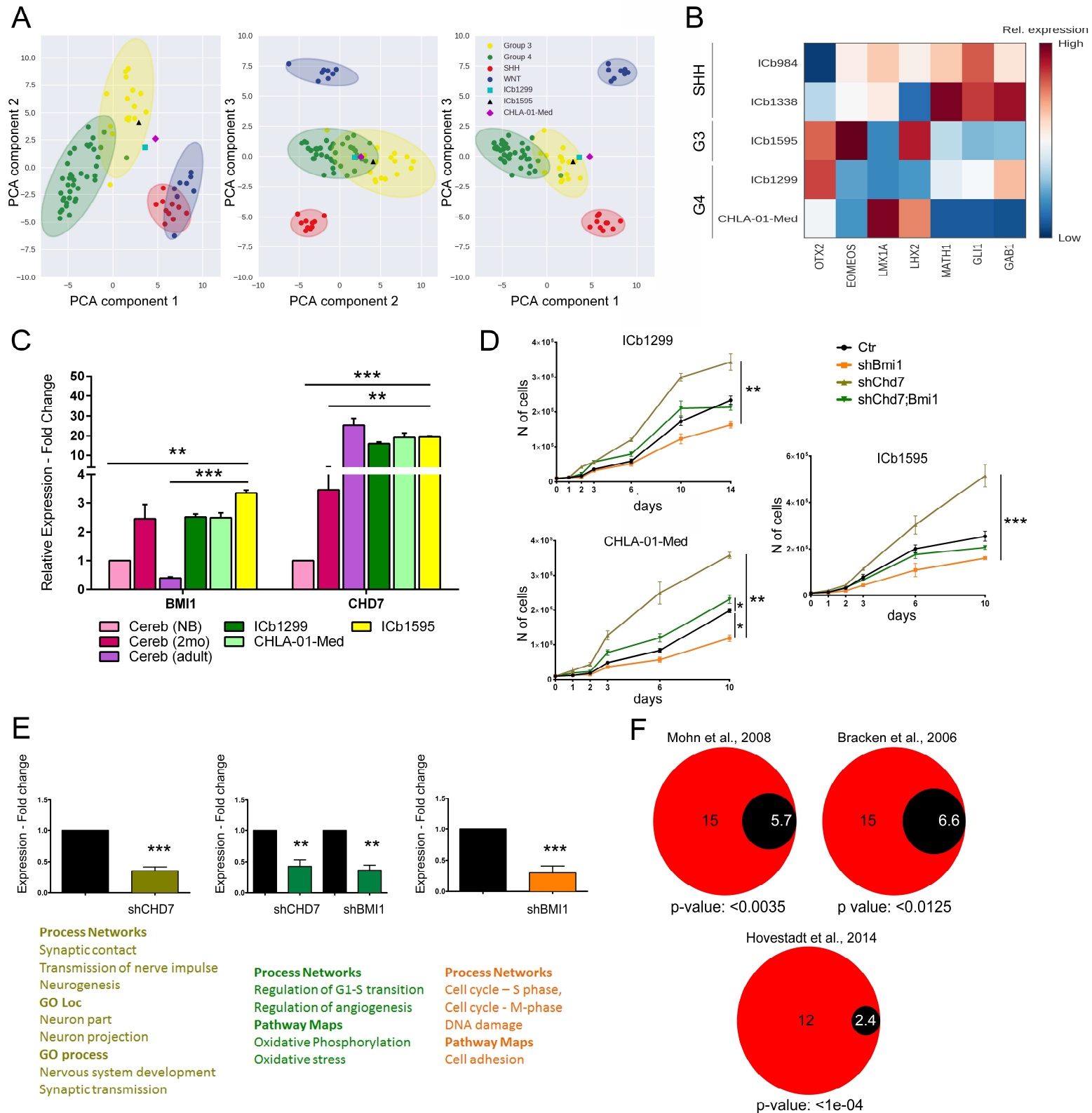
BMI1 High *CHD7* Low vs Rest – p=0.03028



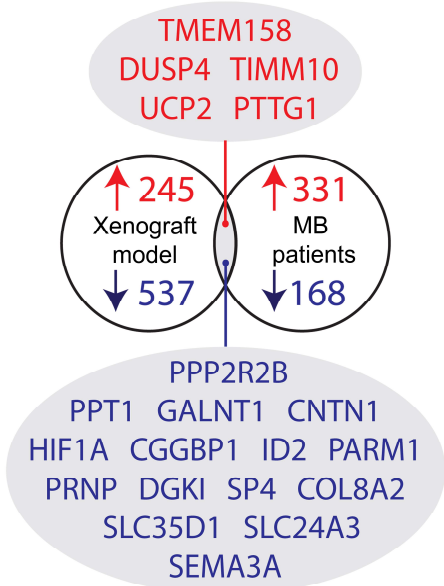
E

Group 4 BMI1 High *CHD7* Low vs Rest – p=0.00154

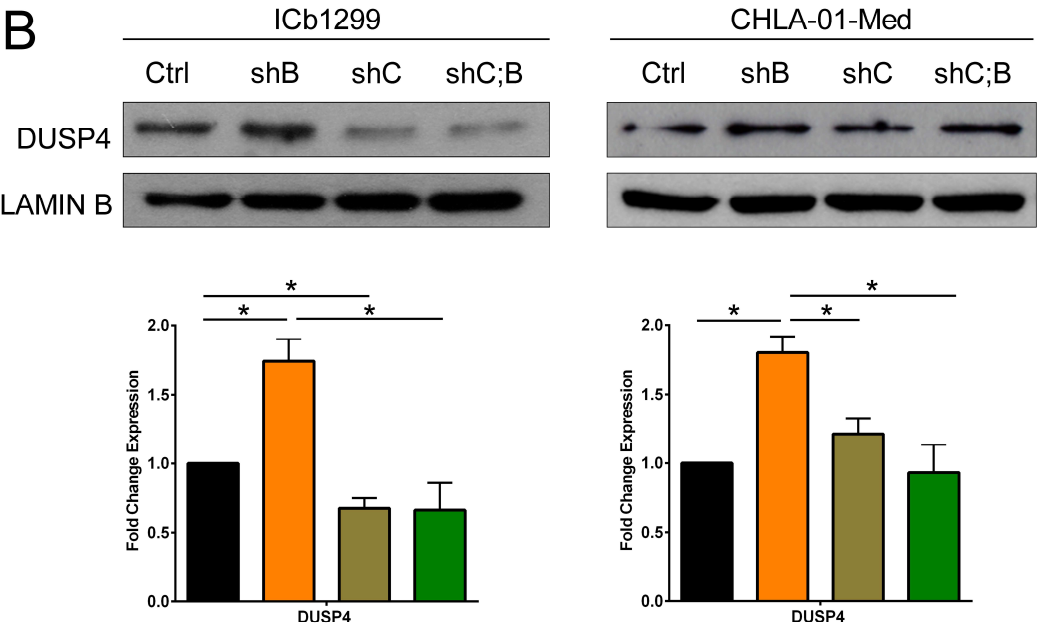




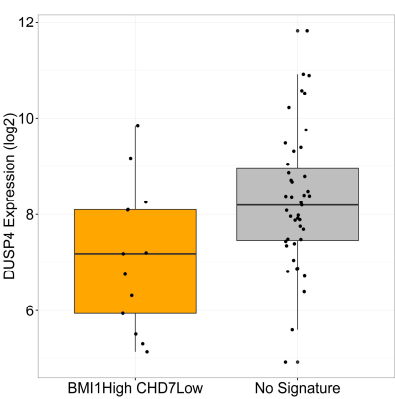
A



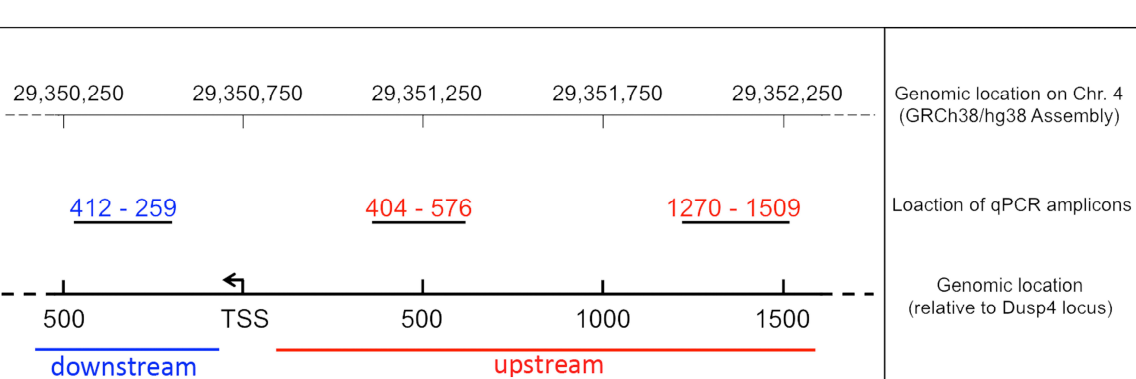
B



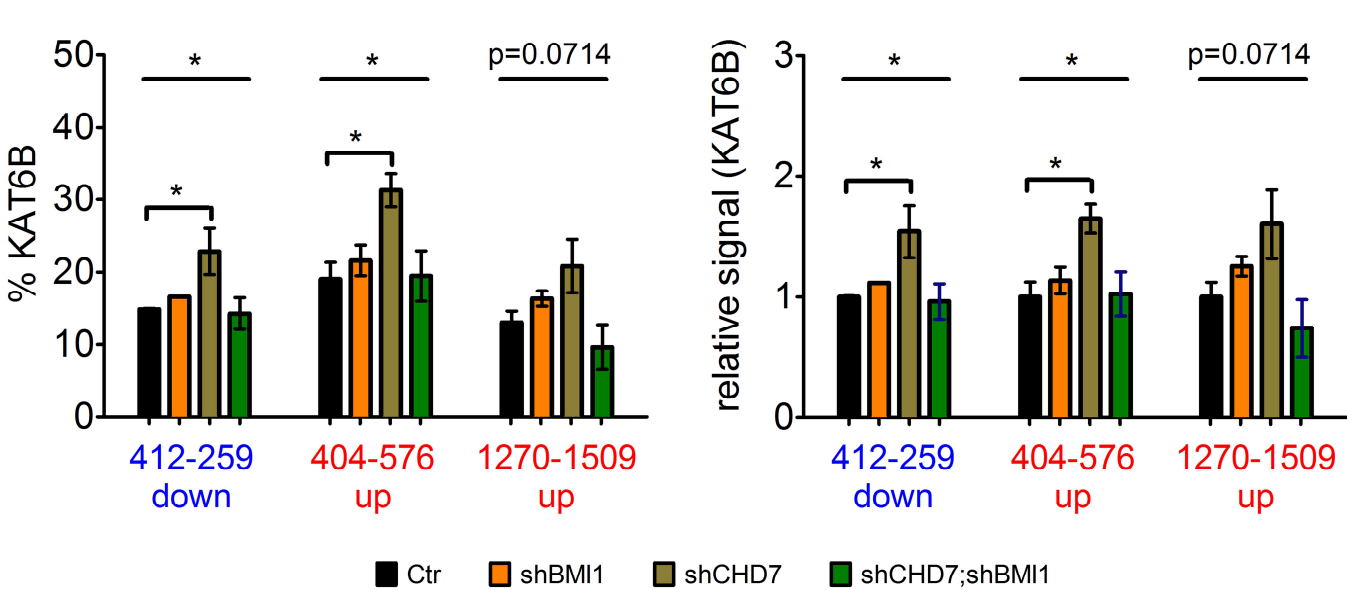
C

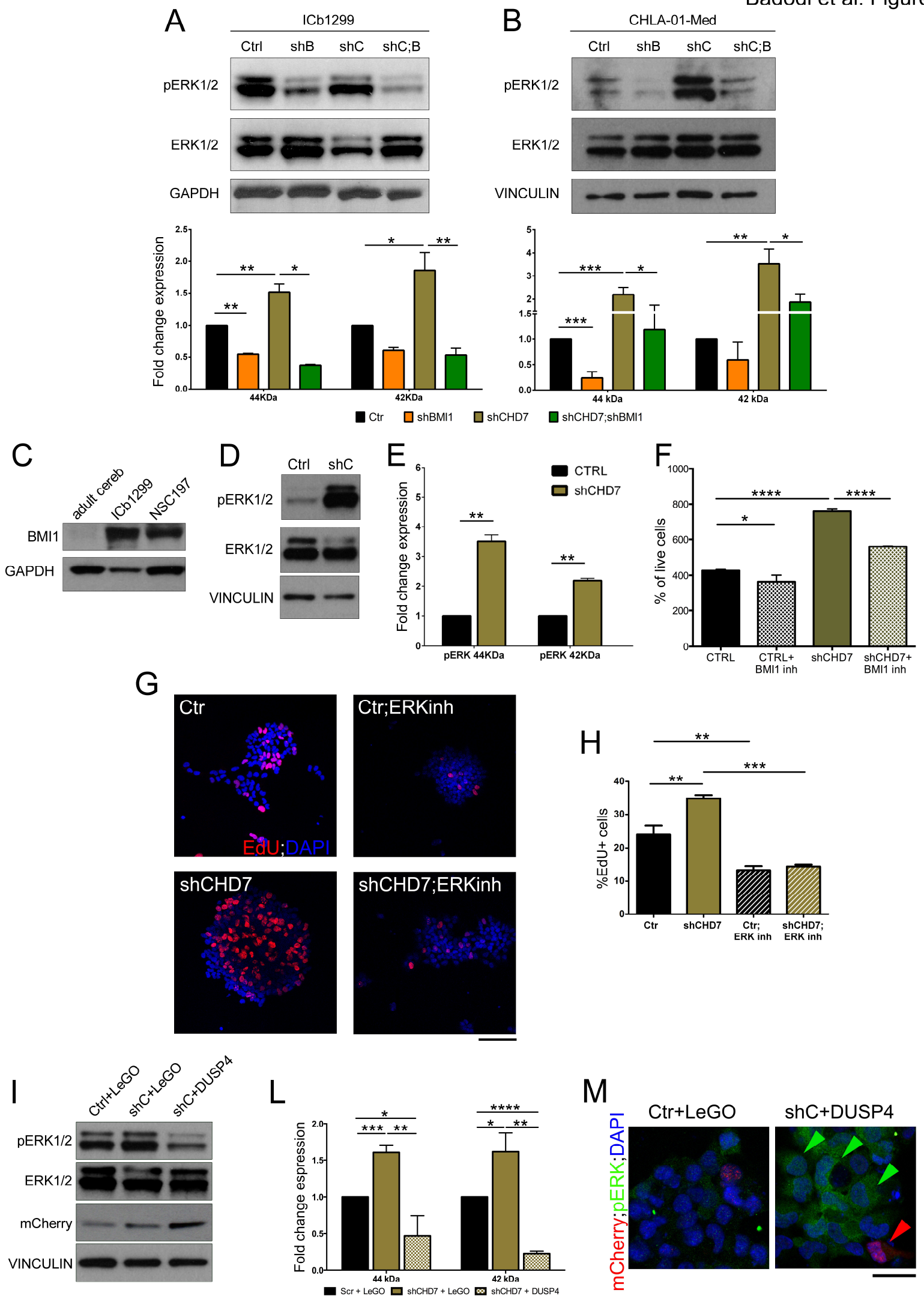


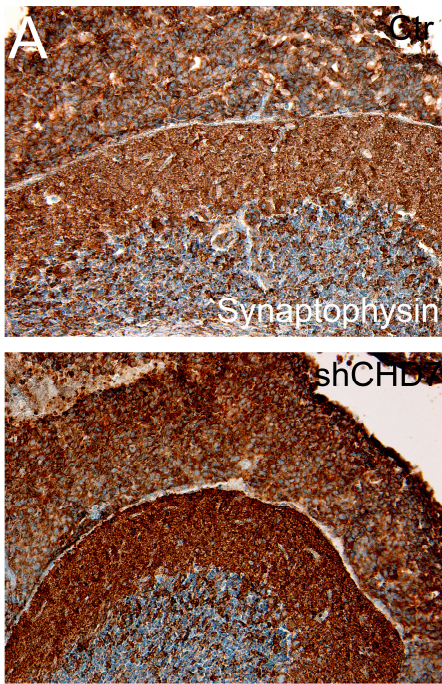
D



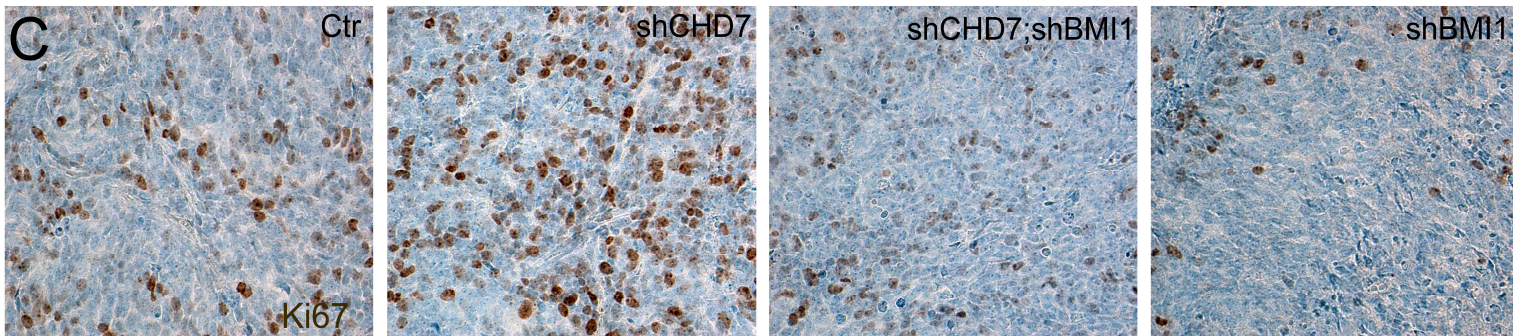
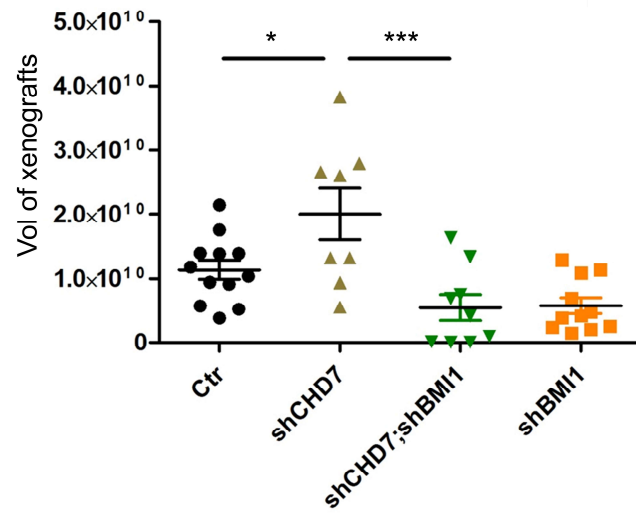
E



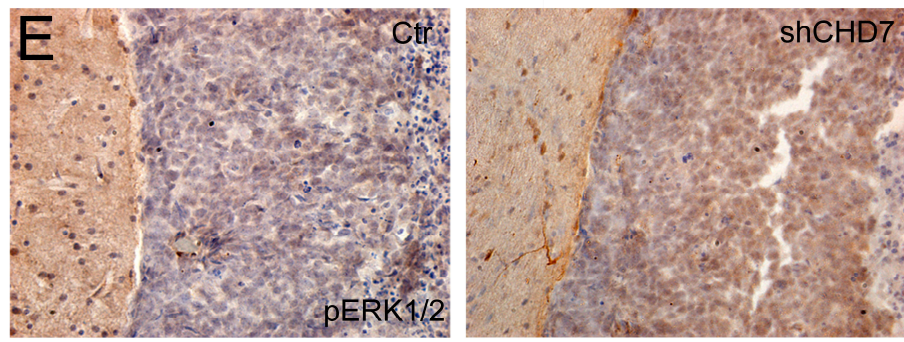
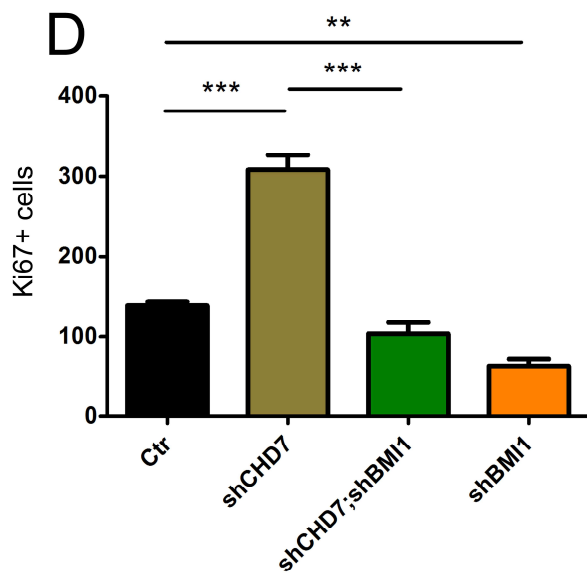


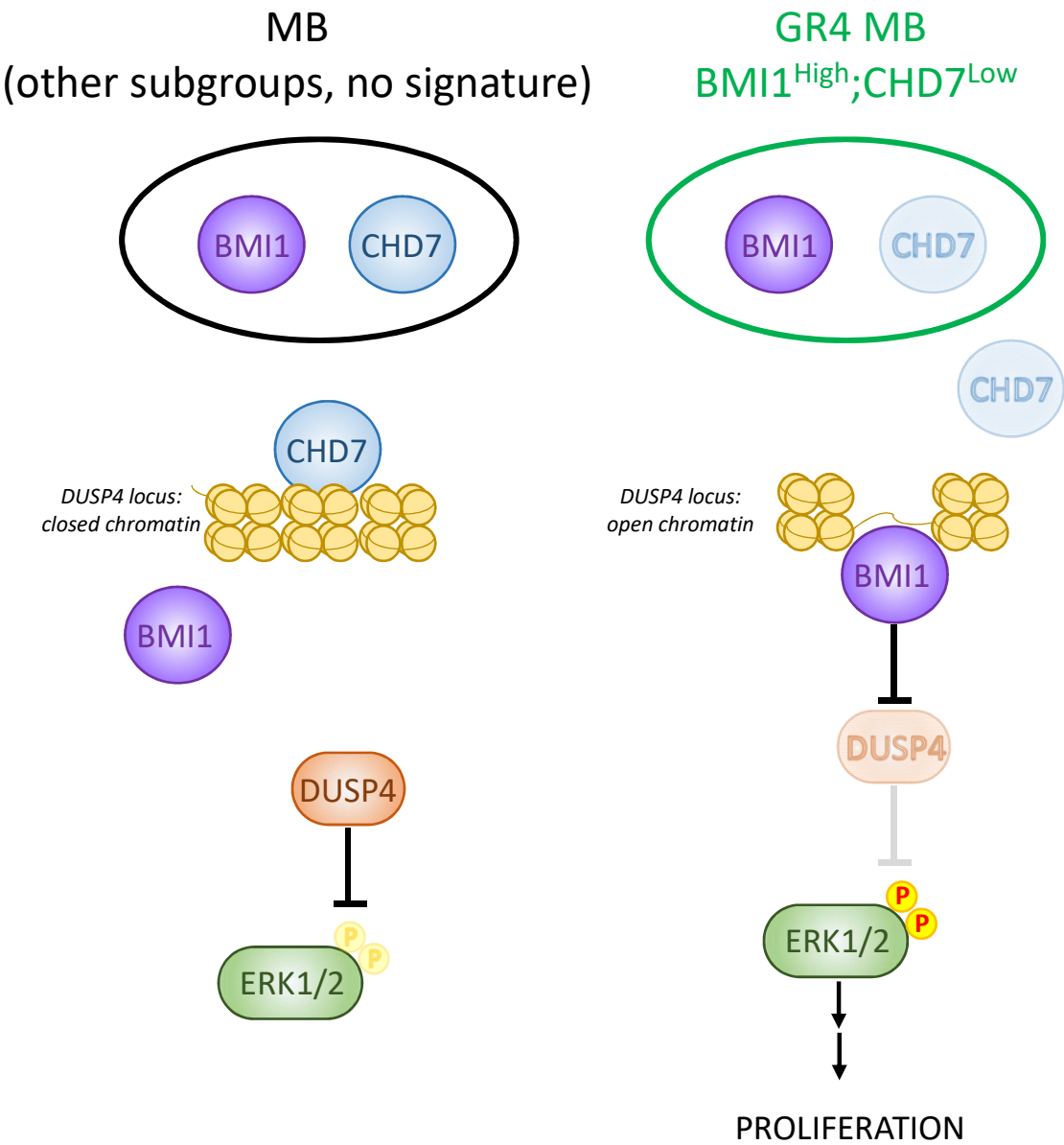


B

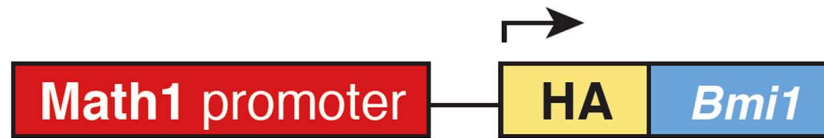


D

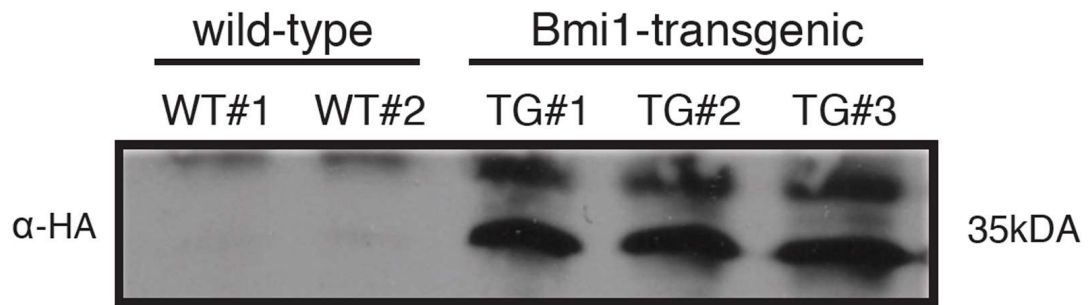




A



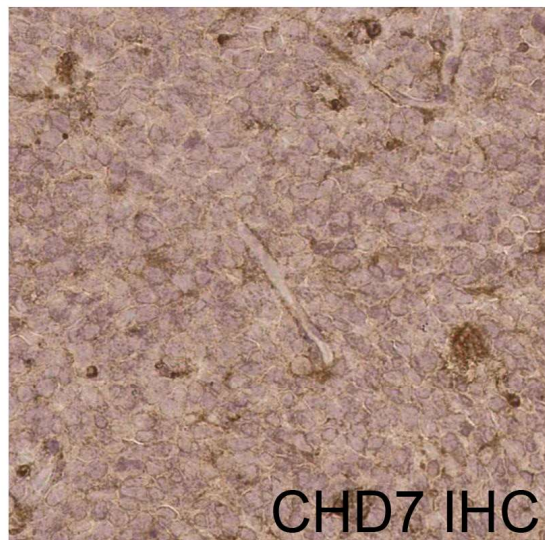
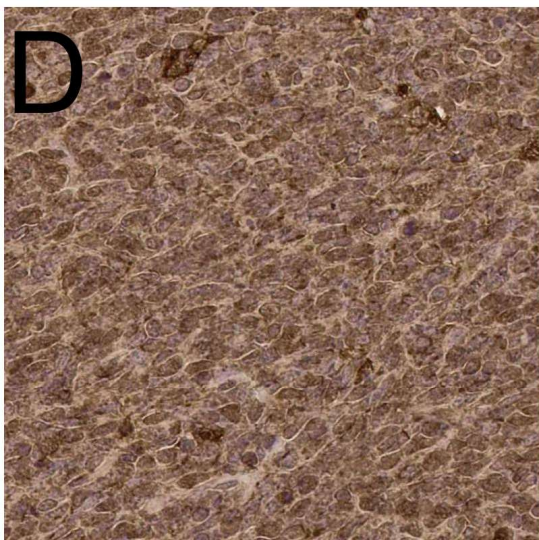
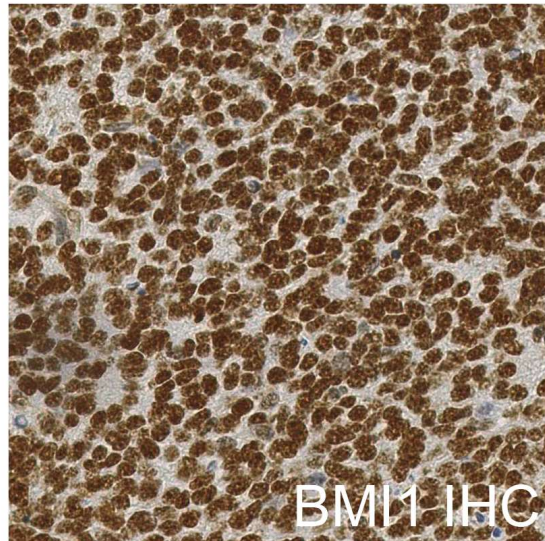
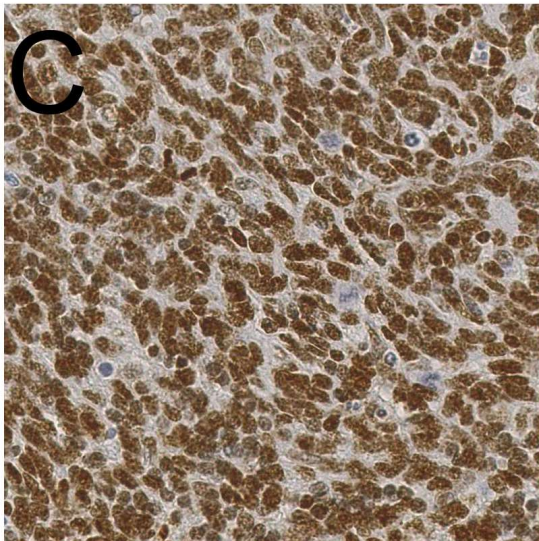
B

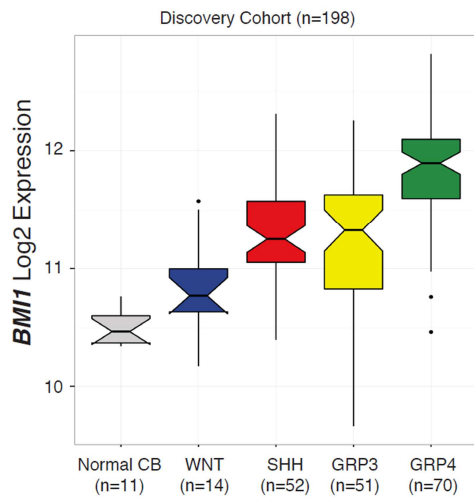
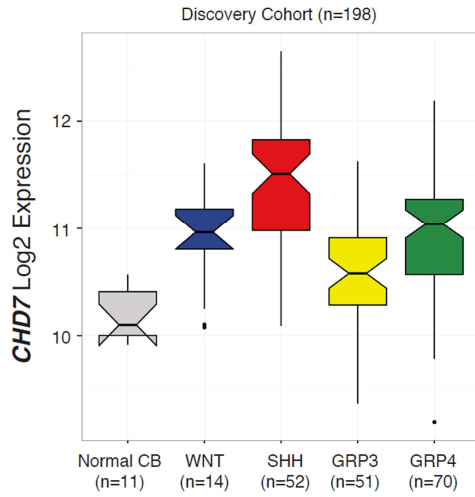


Math1-BMI1;SB11;T2Onc2

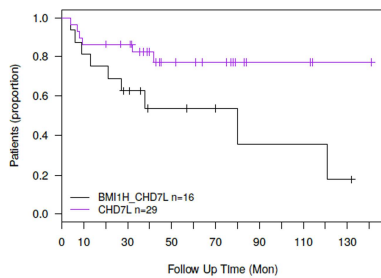
No CHD7 insertion

CHD7 insertion

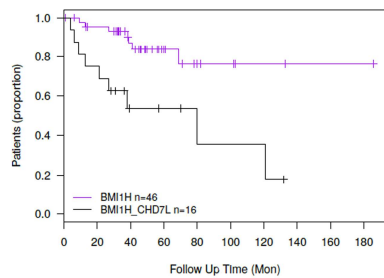


A**B****C**

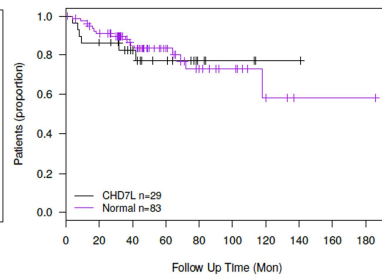
BMI1 High CHD7 Low vs CHD7 Low - $p=0.04516$



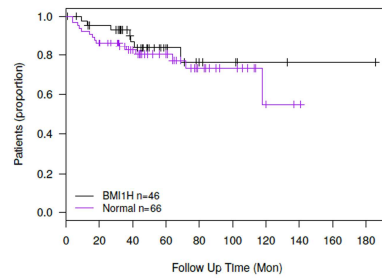
BMI1 High CHD7 Low vs BMI1 High - $p=0.00485$



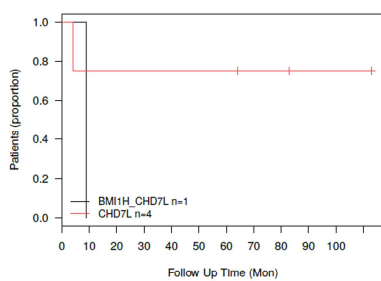
CHD7 Low vs others - $p=0.83576$



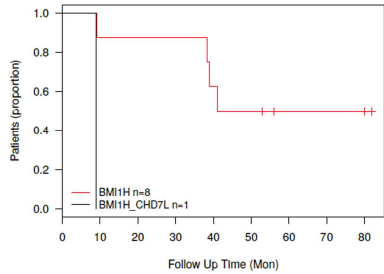
BMI1 High vs others - $p=0.43972$



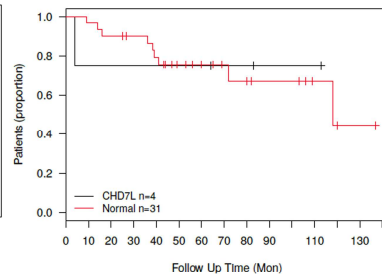
SHH BMI1 High CHD7 Low vs CHD7 Low - $p=0.39495$



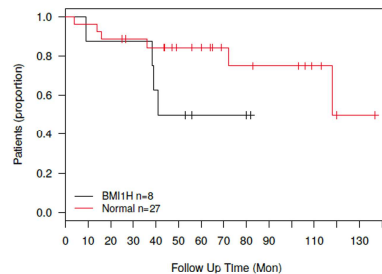
SHH BMI1 High CHD7 Low vs BMI1 High - $p=0.03606$



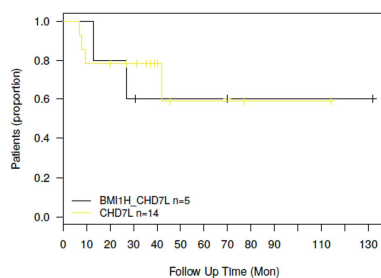
SHH CHD7 Low vs others - $p=0.97097$



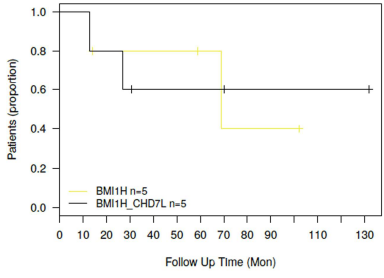
SSH BMI1 High vs others - $p=0.14958$



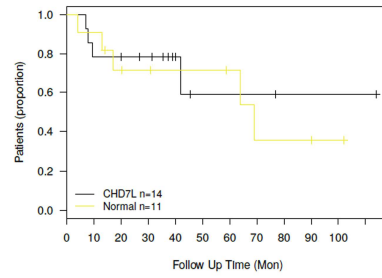
Group 3 BMI1 High CHD7 Low vs CHD7 Low - $p=0.8517$



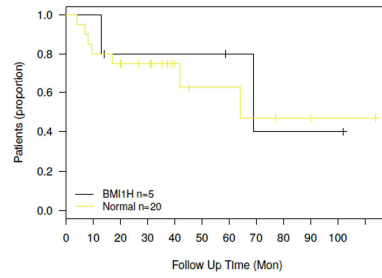
Group 3 BMI1 High CHD7 Low vs BMI1 High - $p=0.9429$



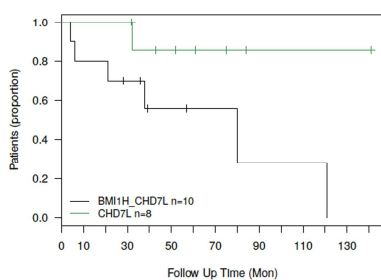
Group 3 CHD7 Low vs others - $p=0.71748$



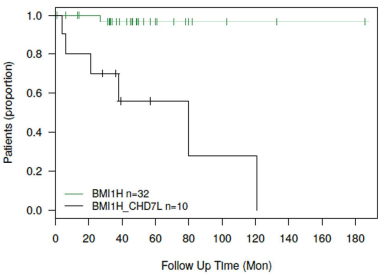
Group 3 BMI1 High vs others - $p=0.80405$



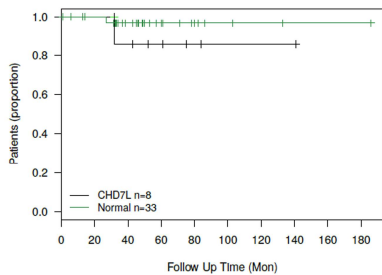
Group 4 BMI1 High CHD7 Low vs CHD7 Low - $p=0.0461$



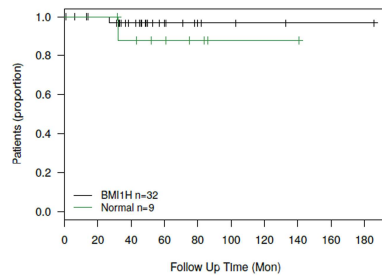
Group 4 BMI1 High CHD7 Low vs BMI1 High - $p=0.0005$

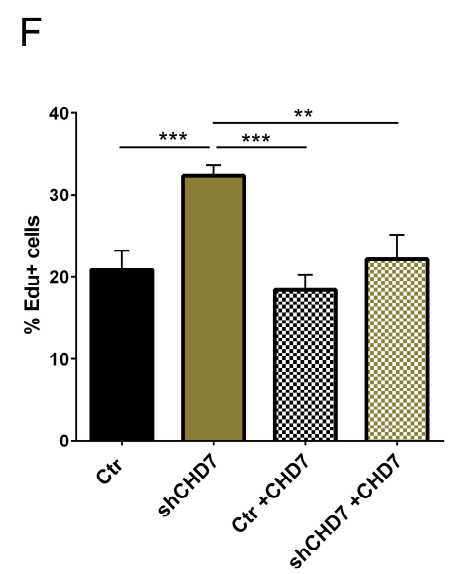
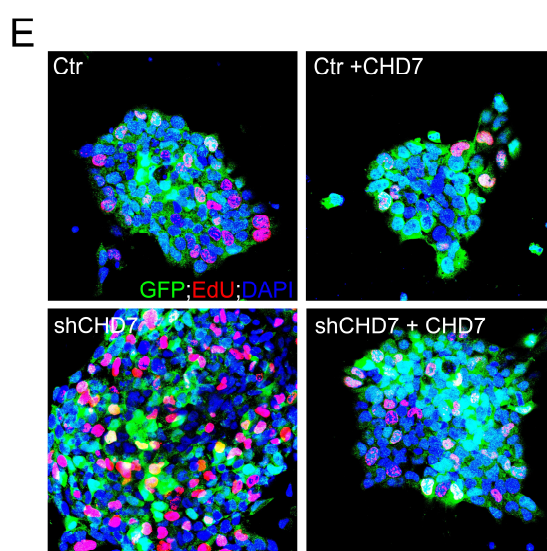
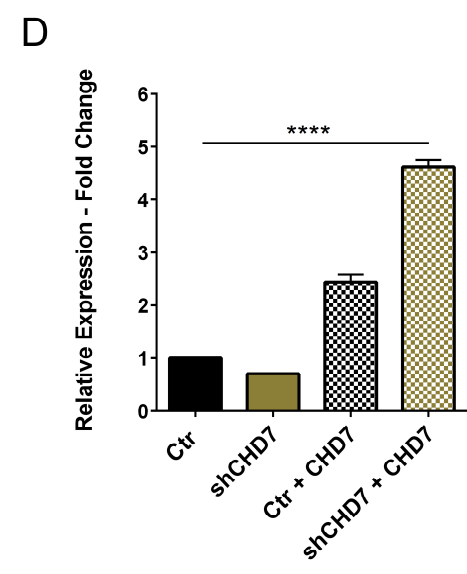
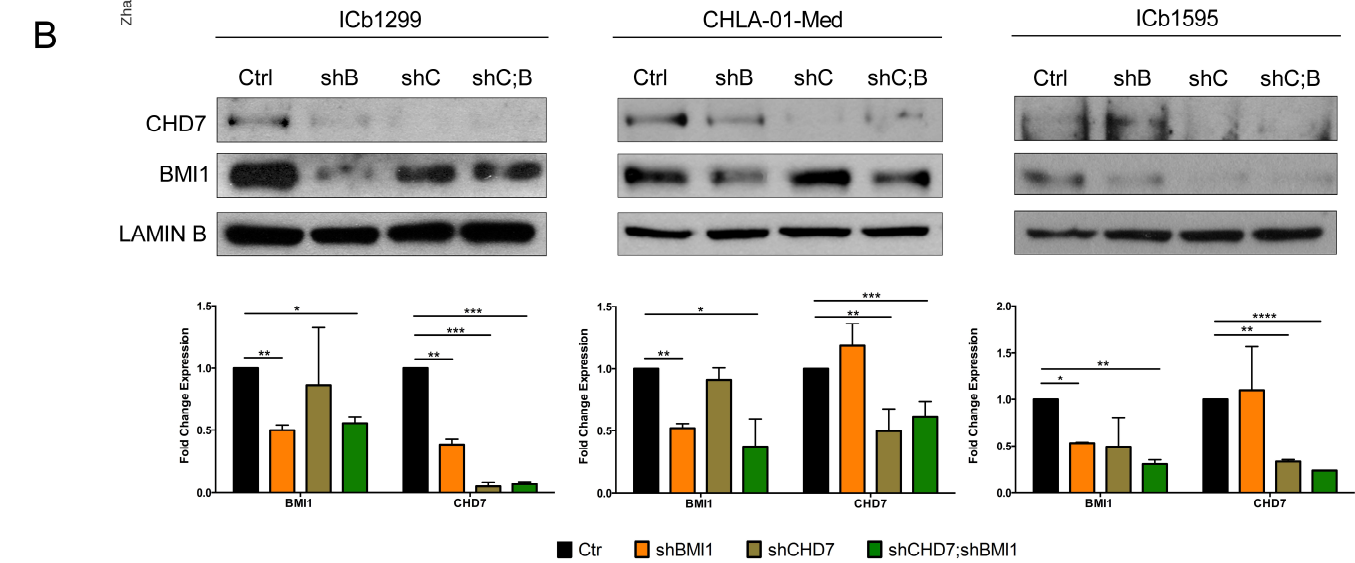
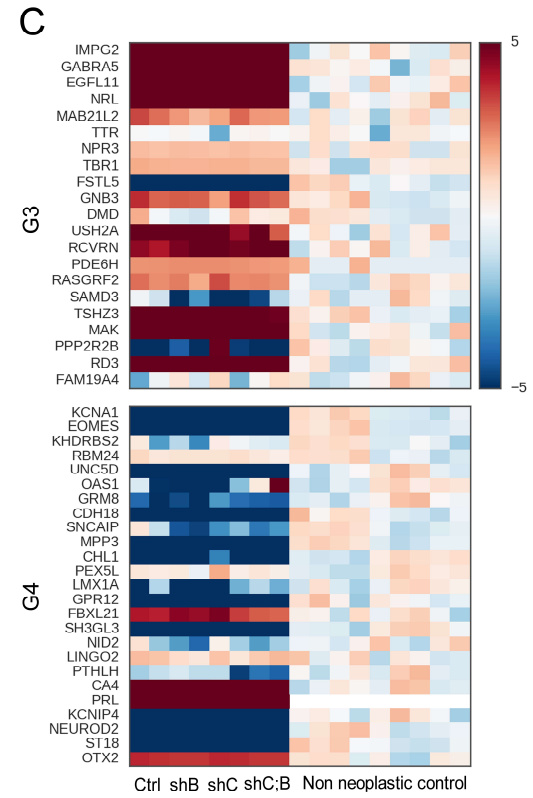
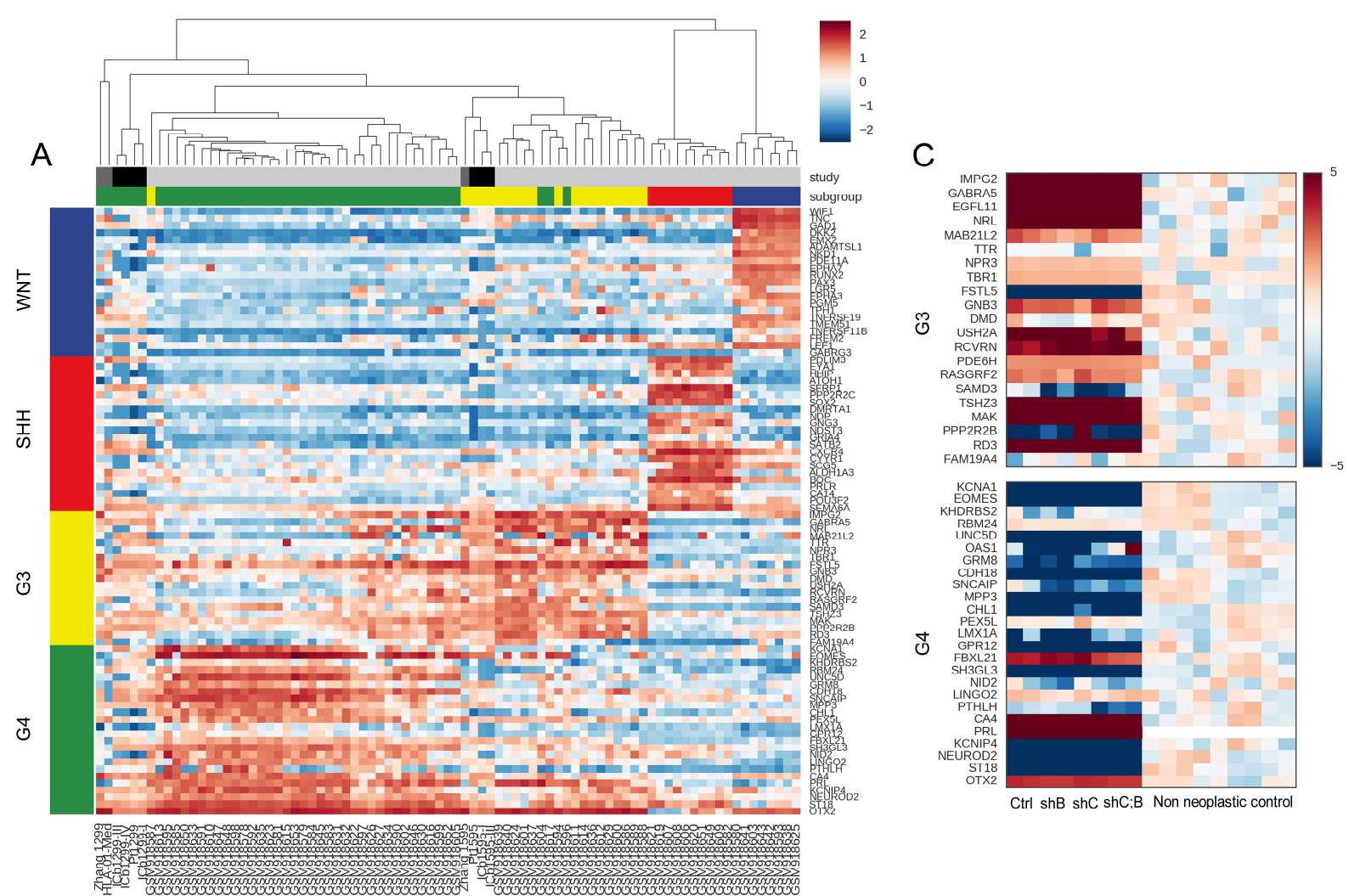


Group 4 CHD7 Low vs others - $p=0.36723$



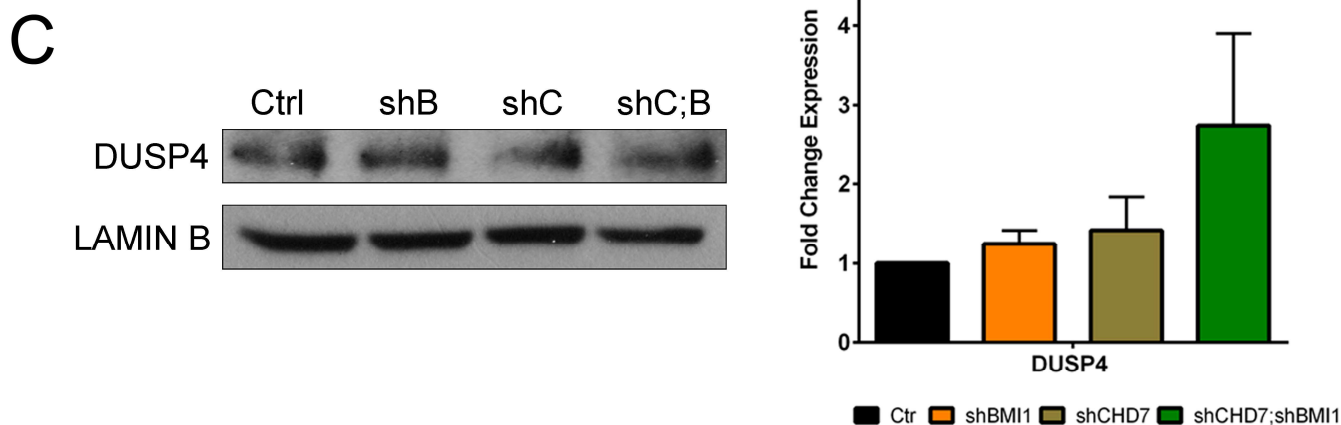
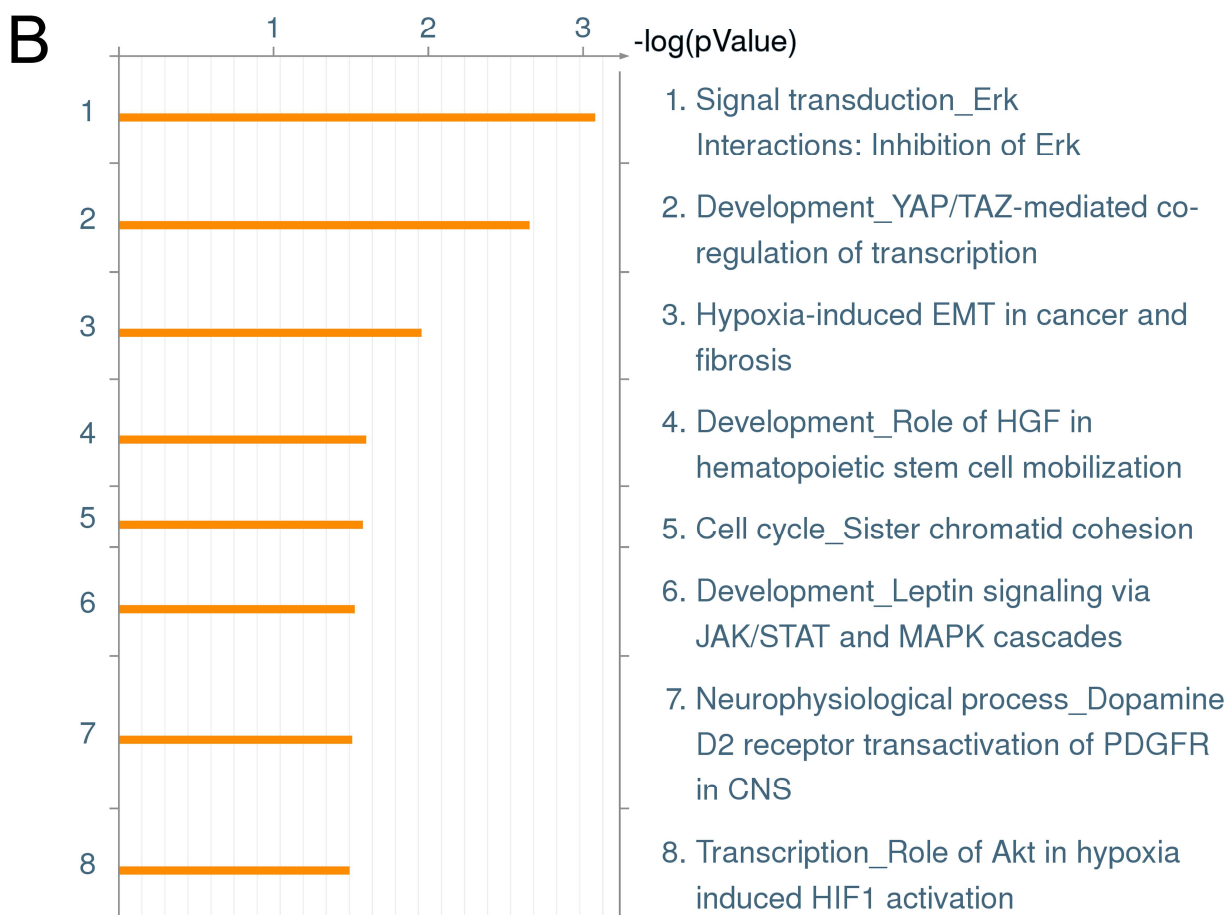
Group 4 BMI1 High vs others - $p=0.42561$



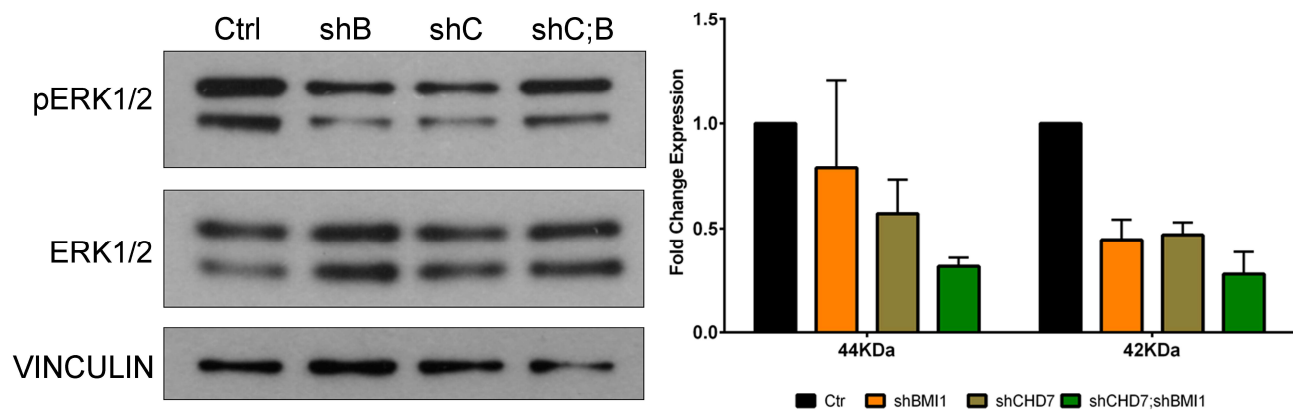


A

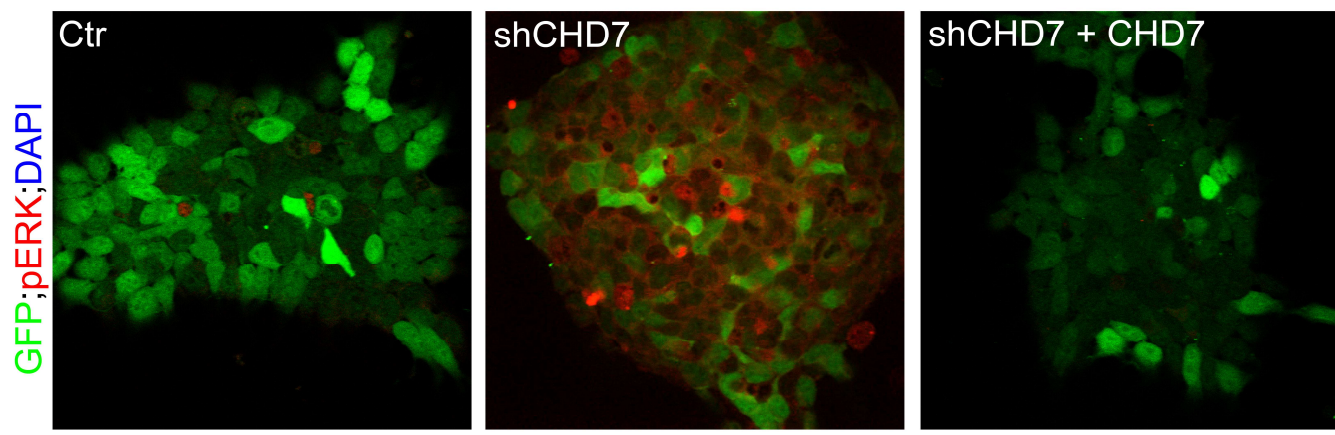
gene	shCHD7	shBMI1shCHD7	lcb1299_FC	lcb1299_p_value	lcb1299_q_value	MB_FC_ratio	MB_p_value	MB_sig	MB_sig	Status
SLC35D1	3.00135	1.96715	-1.525735201	0.04215	0.125181	-1.078730405	0	yes	yes	DOWN
DGKI	3.57766	1.08168	-3.307503143	0.0237	0.10338	-1.042759699	0.004	yes	yes	DOWN
SLC24A3	5.85064	3.49868	-1.672242103	0.0182	0.0994977	-1.130926405	0.004	yes	yes	DOWN
ID2	15.3534	10.0514	-1.527488708	0.0451	0.13191	-1.128258979	0.004	yes	yes	DOWN
GALNT1	19.5259	11.5071	-1.696856723	0.0023	0.0437563	-1.035042798	0.007	yes	yes	DOWN
CNTN1	39.1922	27.1748	-1.442225886	0.01705	0.0947374	-1.248564472	0.007	yes	yes	DOWN
PARM1	11.489	7.49711	-1.532457173	0.01285	0.0833449	-1.116507523	0.011	yes	yes	DOWN
HIF1A	13.0248	7.27331	-1.790766515	0.0048	0.0585563	-1.018513181	0.026	yes	yes	DOWN
SP4	2.95444	1.81172	-1.630737642	0.02205	0.100642	-1.052829404	0.028	yes	yes	DOWN
PPT1	46.3811	31.2354	-1.48488894	0.0115	0.0769811	-1.03633594	0.033	yes	yes	DOWN
PPP2R2B	33.5442	21.4357	-1.564875418	0.0143	0.0851283	-1.11565787	0.04	yes	yes	DOWN
SEMA3A	2.4877	1.03614	-2.400930376	0.00065	0.0342897	-1.230436811	0.041	yes	yes	DOWN
PRNP	85.2033	59.8327	-1.424025658	0.021	0.100642	-1.029101168	0.042	yes	yes	DOWN
COL8A2	1.897	1.11297	-1.704448458	0.047	0.136725	-1.038633848	0.045	yes	yes	DOWN
CGGBP1	29.3201	19.3095	-1.518428753	0.01535	0.0876301	-1.072158891	0.049	yes	yes	DOWN
TMEM158	17.167	25.0201	1.457453253	0.0384	0.122867	1.168853411	0	yes	yes	UP
TIMM10	76.3064	129.571	1.698035814	0.0022	0.0437563	1.048232718	0.021	yes	yes	UP
UCP2	19.7778	29.1342	1.473075873	0.03025	0.115224	1.071493335	0.037	yes	yes	UP
DUSP4	5.55553	8.17981	1.472372573	0.0265	0.108734	1.141653888	0.041	yes	yes	UP
PTTG1	126.979	191.479	1.507958009	0.01025	0.0761169	1.070075107	0.048	yes	yes	UP



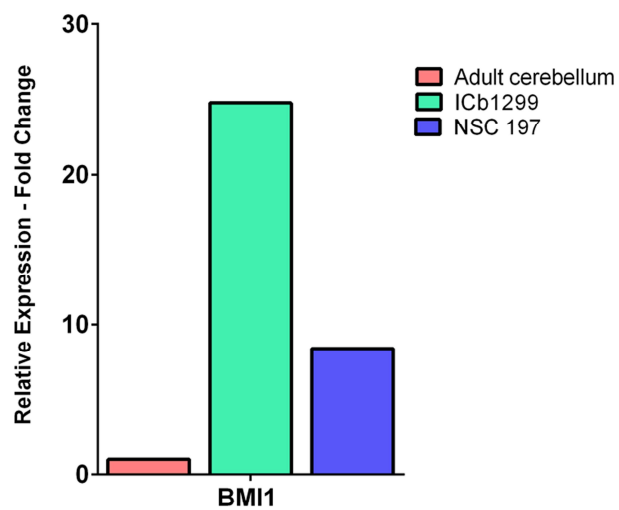
A



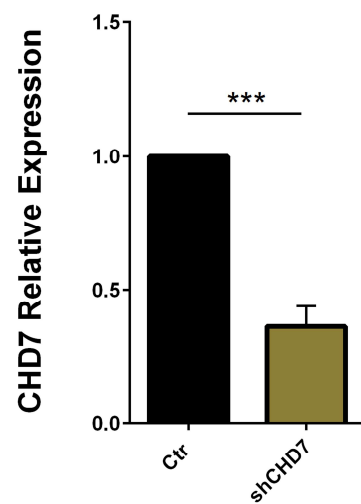
B



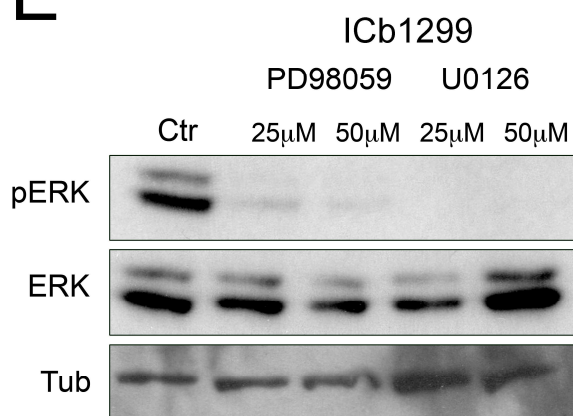
C



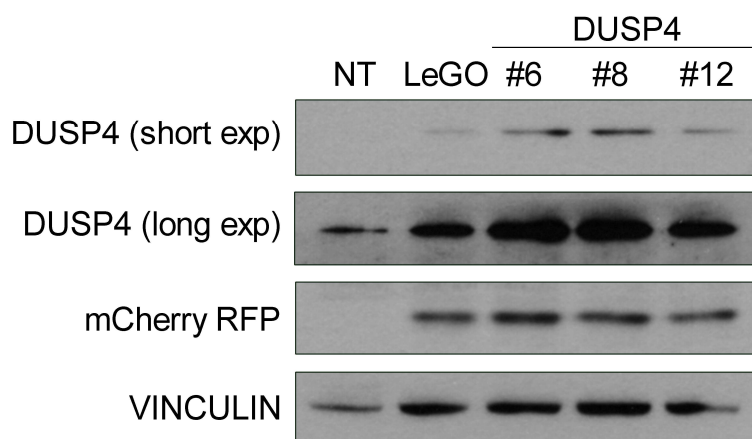
D

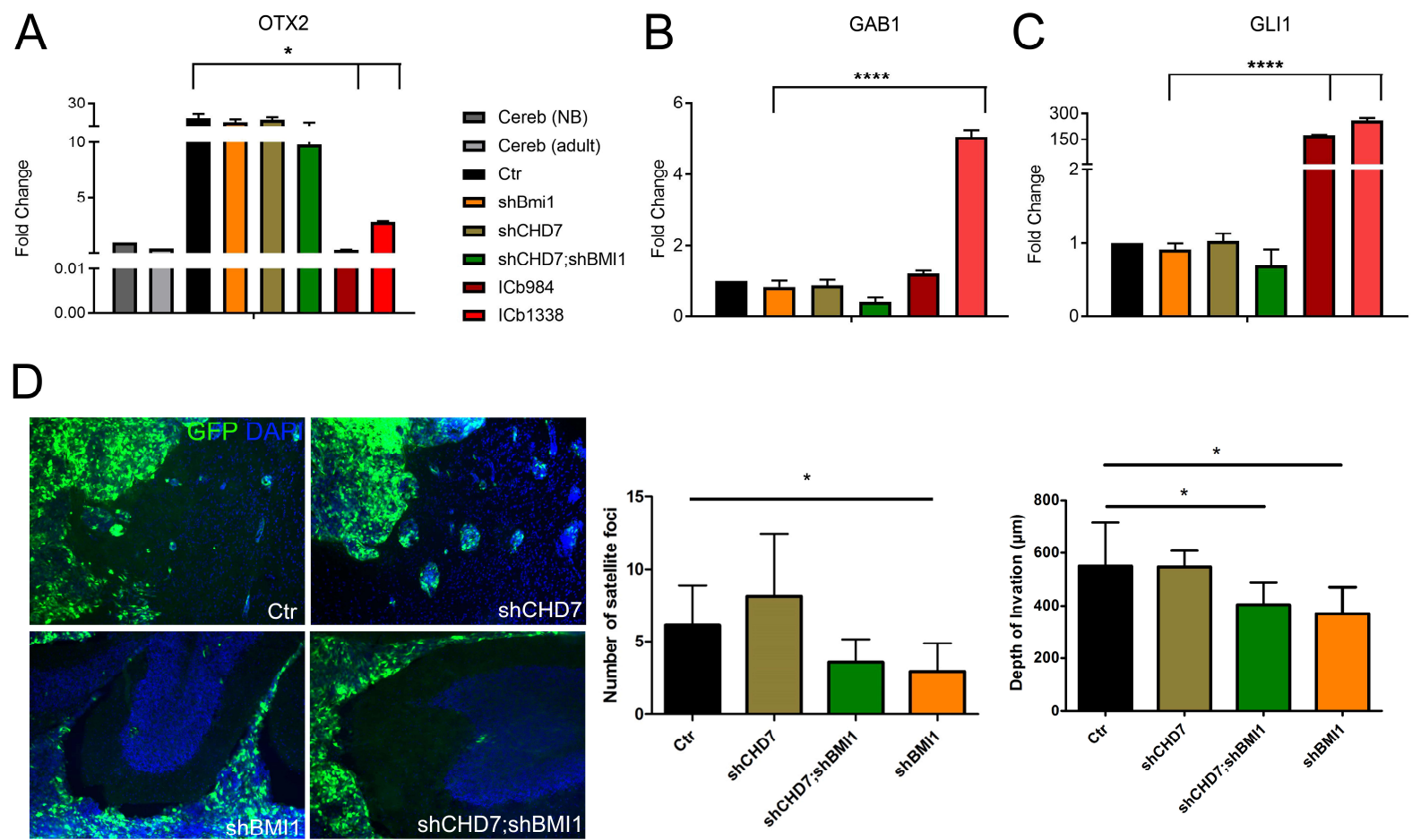


E



F





Supplementary Figure Legends

Figure S1: **A:** Construct of genomic sequence used to generate Bmi1 transgenic/overexpressing animal model under the MATH1-promoter/enhancer. **B:** Anti-HA western blot validating the overexpression of Bmi1 in transgenic (TG) lines as compared to wild-type (WT) littermates. **C:** Expression of Bmi1 in SB MB. **D:** Low expression of Chd7 in SB MB with inactivating insertion in the Chd7 locus. Scale bar is 125 μm .

Figure S2: **A-B:** Box plot representation of BMI1 expression across a cohort of 198 primary MB profiled on Affymetrix 1.1ST gene arrays. BMI1 was significantly overexpressed in G4 MBs versus other molecular subgroups. Conversely, CHD7 was significantly downregulated in G4 versus other molecular subgroups. **C:** Kaplan Meier survival analysis of patients with the BMI^{High};CHD7^{Low} signature relative to BMI^{High} only and CHD7^{Low} only. Across the full cohort, the BMI^{High};CHD7^{Low} signature is associated with significantly poorer outlook compared to the single phenotypes. Dividing the cohort into subgroups demonstrates that this finding is consistently reproduced only in G4.

Figure S3 A: Heatmap showing relative gene expression for 100 medulloblastoma signature genes (Lin et al., 2016) for the Robinson cohort (light grey), tumour and early passage xenograft culture data from Zhao et al. (black; tumour samples have names beginning Pt) and the current study (dark grey). The dashed line indicates that 5 clusters are detected. The Robinson data are clustered into the expected subgroups with three errors. ICb1595 cluster with other G3 samples. ICb1299 and CHLA-01-Med cluster separately but closer to G3 and G4 than WNT and SHH. **B:** Expression levels of BMI1 and CHD7 upon silencing of BMI1 or

CHD7 or both ($n = 3$, \pm SEM). **C:** Log2 gene expression levels for 50 signature genes commonly associated with G3 and G4 medulloblastoma in the single, double and control knockdown conditions of line ICb1299 (in duplicate), relative to healthy cerebellum tissue. The ICb1299 samples show a consistent profile across all conditions. **D:** Reconstitution of CHD7 expression is shown by qRT-PCR in ICb1299 G4 MB where CHD7 had been previously silenced. **E:** Increased proliferation as assessed by EdU pulse/chase upon CHD7 silencing is rescued upon reconstitution of CHD7 expression in ICb1299. **F:** Quantification of the findings. Scale bar is 50 μ m.

Figure S4: **A:** Summary table of DE genes in G4 MB cells upon BMI1 knockdown in a CHD7 silenced context and in G4 MB with a BMI^{High};CHD7^{Low} signature. **B:** Metacore analysis of the 20 DE genes shows an impact on ERK signalling. **C:** Western blot and quantification of DUSP4 expression show that BMI1 doesn't regulate DUSP4 in G3 MB cells ($n = 3$, \pm SEM).

Figure S5: **A:** No overactivation upon CHD7 silencing in G3 MB cells. ($n=2$, \pm SEM). **B:** Immunofluorescence analysis showing rescue of the ERK pathway overactivation upon CHD7 reconstitution in ICb1299. Scale bar is 50 μ m. **C:** Quantification of Western blot analysis in Fig. 5C shows overexpression of BMI1 in human NPC as compared to adult cerebellum. **D:** Efficient silencing of CHD7 in the cells analysed in Fig. 5D and E. **E:** Western Blot analysis showing effective pharmacological inhibition of ERK1/2 pathway in ICb1299. **F:** Western blot analysis of HEK293 cells overexpressing 3 different LeGO-mCherry-DUSP4 clones (#6, 8 and 12) shows increased levels of both mCherry and DUSP4 expression.

Figure S6: A: OTX2 expression is unchanged upon CHD7 and/or BMI1 silencing in G4 MB xenografts and is higher than normal cerebellum and SHH MB cells (ICb984 and ICb1338).

B-C: GAB1 and GLI1 expression in G4 MB xenografts is lower than SHH MB cells. **D:**

Intraparenchymal invasion of G4 MB cells, as assessed by quantification of the number of satellite foci around the main tumour bulk (left) or as depth of invasion (right) is dependent on BMI1 but not on CHD7. Data are represented as mean \pm SEM. Scale bar is 125 μ m.

References:

Lin, C. Y., Erkek, S., Tong, Y., Yin, L., Federation, A. J., Zapatka, M., Haldipur, P., Kawauchi, D., Risch, T., Warnatz, H. J., *et al.* (2016). Active medulloblastoma enhancers reveal subgroup-specific cellular origins. *Nature* 530, 57-62.

Extended Experimental procedures

Analysis of gCIS

A description of the linker-mediated PCR and Illumina HiSeq sequencing used for transposon insertion sites as well as details of the gCIS prediction, of the genetic algorithm for driver gene prediction and of the PCR for Sleeping Beauty tagged fragments are presented in (Morrissy et al., 2016). Briefly, sequenced libraries were demultiplexed and aligned as described (Brett et al., 2011). Custom scripts were used to demultiplex and trim SB transposon sequences, and reads were aligned to mouse assembly NCBI37/mm9 (July 2007) using Novoalign. Statistical enrichment of transposon integration events within each of 19,000 mouse RefSeq genes was assessed using a chi-squared test considering the following: the number of TA dinucleotide sites within the gene relative to the number of TA sites in the genome, the number of integration sites within each tumour, and the total number of tumours in each cohort. The p-value calculated for each of gene was adjusted for multiple hypothesis testing using Bonferroni correction. Finally, significant gCIS predictions were manually curated to filter out artefacts and local hopping events. The excel files of this analysis and of the comparative analysis with a published dataset of SB-induced insertion sites in MB occurring in *Ptch1*^{+/-} mice (Morrissy et al., 2016) are included as supplementary material (*Bmi1* filtered.clonal gCIS_3 and *Bmi1*-SB versus SB and *Ptc*-SB analysis-August 13-12).

DUSP4 lentiviral construct cloning

DUSP4 coding sequence was PCR cloned from plasmid R777-E039 Hs.DUSP4 (Addgene, 70323) and inserted in the lentiviral vector LeGO-iC carrying the fluorescent reporter RFP-mCherry with Gibson Assembly cloning strategy (NEB). The resulting *LeGO-iC-DUSP4-mCherry* was sequence verified. **Production of shRNA lentiviral vector and gene silencing**

GIPZ lentiviral shRNA vectors containing a hairpin sequence targeting BMI1 or CHD7 and the coding sequences for GFP and puromycin resistance gene were purchased from Dharmacon, UK. Packaging, virus production and determination of titre were carried out as previously reported (Merve et al., 2014). Cells were infected overnight at a multiplicity of infection (MOI) of 12. After 96h from the infection, puromycin selection at a concentration of 2.5µg/ml was applied to enrich for the transduced population. The efficacy of gene silencing was assessed by qRT-PCR or Western blot analysis. All experiments were conducted at least in triplicates.

Growth Curve

Transduced patient-derived cells (Scramble control, shBMI1, shCHD7, shCHD7;shBMI1) were seeded in 24-well plates at the same density. For the following 10 days, cells were harvested at specific time points (1, 2, 3, 6 and 10 days) and the number of proliferating cells was counted with a hemocytometer and Trypan Blue staining. Experiments were carried out at least two times in triplicate.

EdU staining

The Click-iT® EdU Alexa Fluor® 594 Imaging Kit (ThermoFisher) was used to assess EdU incorporation following manufacturer recommendations. Briefly, transduced cells were seeded on coated coverslips in triplicates, treated with 10µM of EdU solution for 3 hours and then fixed in 4% paraformaldehyde (PFA). Cells were then washed, blocked and incubated with Click-iT® reaction cocktails containing Alexa Fluor® azide. Rabbit polyclonal GFP antibody (1:1000, ab290, abcam) was used to visualize transduced cells. Five representative images (40X magnification) of each sample were captured using a Zeiss 710 Confocal Microscope. The percentage of positive nuclei was calculated as the ratio between EdU-positive cells and the total number of nuclei counted using ImageJ software.

Human neural stem cells (NSCs) culture

Human fetal brains were obtained from consenting patients, as approved by Hamilton Health Sciences/McMaster Research Ethics Board. Samples were dissociated in PBS containing 0.2 Wunisch unit/mL Liberase Blendzyme (Roche), and incubated at 37°C in a shaker for 15 min. The dissociated tissue was then filtered through 70µm cell strainer and collected by centrifugation at 1500 rpm for 3 min. Neural stem cells were resuspended in Neurocult complete media, (STEMCELL Technologies), supplemented with human recombinant epidermal growth factor (20ng/mL: STEMCELL Technologies), basic fibroblast growth factor (20ng/mL; STEMCELL Technologies), antibiotic-antimycotic (10mg/mL; Wisent), and plated in ultra-low attachment plates (Corning). Red cells were lysed using ammonium chloride solution (STEMCELL Technologies). Neural stem cells were grown in suspension in Neurocult complete for 2 days. Thereafter, the neurospheres were plated onto polyornithine-laminin (PL) coated tissue culture treated plates. Once propagated, the cells were dissociated with TrypLE (Invitrogen) and replated onto appropriately sized plates for subsequent experimentation.

NSCs transduction and propagation

NSCs were plated onto 24 well PL plates at a density of 2.0×10^5 cells per well for transduction with appropriate viruses. The cells were incubated in 37°C for 24 hours and the media was changed completely. Thereafter, the cells were incubated further in 37°C for 48 hours. The transduced cells were selected with the addition of 1µg/mL puromycin in Neurocult complete media. The remaining cells after puromycin selection, were then propagated further and replated if necessary. Transduced NSC were dissociated and adjusted to 1 million single cells/mL in PBS+2mM EDTA. Samples were analysed and/or sorted using a MoFlo XDP cell sorter and Kaluza software (Beckman Coulter). Dead cells

excluded using the viability dye 7-AAD (1:10; Beckman Coulter). Cells were sorted into Neurocult complete media and allowed to equilibrate at 37°C overnight prior to use in experiments and replating onto PL plates.

RNA extraction and qRT-PCR analysis

Total RNA was isolated from cell pellets with RNeasy Micro purification kit (Qiagen) and digested with DNaseI (Applied Biosystems). 8 to 12 xenograft frozen sections were used to extract total RNA by the proteinase K/acid phenol method (Khodosevich et al., 2007) followed by digestion with DNaseI. Total RNA from formalin-fixed, paraffin-embedded (FFPE) samples was isolated with the FFPE RNA/DNA Purification Plus Kit (Norgen Biotek). 4 sections of 20 µm of thickness were used for each sample and DNaseI digestion was performed according to manufacturer instructions. The cDNA synthesis was carried out with SuperScript III Reverse Transcriptase Kit (Invitrogen). Analysis of gene expression was performed with the Applied Biosystems 7500 Real-Time PCR System using TaqMan gene expression MasterMix (Applied Biosystems) and SYBR Green PCR Master Mix (Applied Biosystems) according to standard protocols. Technical triplicates for each samples were analyzed. The Ct values of all the genes analyzed were normalized to the Ct of housekeeping gene (GAPDH for TaqMan probes and average Ct of 18S, ACTB and ATP 5B for SYBR Green PCR) and fold changes were calculated.

Target genes and TaqMan probes used are listed as below:

Target Gene	Assay ID
Bmi1	Hs00180411_m1
CHD7	Hs00215010_m1

EOMEOS	Hs00172872_m1
GAPDH	Hs02758991_g1
GAB1	Hs00157646_m1
GLI1	Hs00171790_m1
LHX2	Hs00180351_m1
LMX1A	Hs00892663_m1
ATOH1 (MATH1)	Hs00245453_s1
OTX2	Hs00222238_m1

Primers used in SYBR Green qPCR are the following:

BMI1 FW: GCTGGTTGCCCATTGACAG; REV: CGATGCATTTCTGCTTGATAA

CHD7 FW: GAAGAAGATATAGAGACCCAC; REV: TCTTTGGTACATAACTTGGC

CHD7_NSC FW: CTTTTCATGAGCCACAAACG; REV: TCTTCTCAAAGCTTTGGTCAC

18S FW: CGCCGCTAGAGGTGAAATTCT; REV: CGAACCTCCGACTTTCGTTCT

ACT B FW: GCGAGAAGATGACCCAGATC; REV: CCAGTGGTACGGCCAGAGG

ATP 5B FW: CCCAGGCTGGTTCAGAGGT; REV: AGGGGCAGGGTCAGTCAAG

GAPDH_NSC FW: TGCACCACCAACTGCTTAGC; REV: GGCATGGACTGTGGTCATGAG

Immunocytochemistry (IHC)

IHC was conducted on FFPE or frozen sections of xenografts tumors (tissue fixed with 4% PFA and embedded in OCT) either manually or with an automated Ventana Discovery XT.

Manual IHC: Briefly, FFPE sections were dewaxed with xylene and hydrated with graded alcohol series. Heat-induced antigen unmasking was performed with Antigen Unmasking Solution, citrate-based (Vector) and endogenous peroxidase activity was quenched with 3% H₂O₂ solution. Sections were treated with 10% Normal Donkey Serum to block non-specific binding sites and incubated overnight with primary antibodies at 4°C. Sections were then incubated with biotinylated secondary antibody (Vector, 1:500) for two hours at room temperature, treated with VECTASTAIN ABC Reagent (Vector) and DAB (SIGMAFAST tablets, Sigma Aldrich). Samples were counterstained with hematoxylin.

The following primary antibodies were used: rabbit polyclonal anti-CHD7 (1:500, ab117522, Abcam), rabbit monoclonal anti-phospho-p44/42 MAPK (1:100, D13.14.4E, Cell Signaling).

Automated IHC: Paraffin embedding, coronal sectioning of 3-µm and staining for haematoxylin and eosin, synaptophysin and human vimentin were performed by UCL IQPath (Institute of Neurology, London, UK). Immunostaining was done on Ventana Discovery XT instrument an automated staining machines (ROCHE, Burgess Hill, UK) following the manufacturer's guidelines, using horseradish-peroxidase-conjugated streptavidin complex and diaminobenzidine as a chromogen. The following antibodies were used for histological characterisation: synaptophysin (Invitrogen 080130, prediluted), KI67 and human vimentin (Roche 790-2917, prediluted).

Immunofluorescence analyses

Immunofluorescence analyses were conducted on Icb1299 cells and frozen tissue sections from xenograft tumours. Icb1299 cells, cultured on Poly-lysine (PLL) coated coverslips, were fixed using 4% PFA. Cells and freshly frozen tissue sections (xenografts fixed with 4% PFA and embedded in OCT) were treated with 10% Normal Goat Serum, followed by incubation

with primary antibodies overnight at 4°C. Appropriate secondary antibodies (1:500, Invitrogen) were used. The samples were counterstained with DAPI. The following primary antibody were used: mouse monoclonal anti-mCherry (1:200, ab125096, abcam); rabbit monoclonal anti-phospho-p44/42 MAPK (1:100, D13.14.4E, Cell Signaling); chicken polyclonal anti-GFP (1:500, ab13970, abcam).

Western blot analysis

Transduced patient-derived cells were lysed for 30 minutes on ice using RIPA lysis buffer supplemented with 2mM PMSF, 1mM sodium orthovanadate and protease inhibitor cocktail (PIC) (RIPA Lysis Buffer System, Santa Cruz) followed by 3 pulses of sonication. Frozen tissue was extensively washed with cold PBS supplemented with PIC to dissolve OCT before lysis in RIPA buffer.

Nuclear and cytoplasmic fractions of transduced cells were obtained with two different lysis buffers. Briefly cells were harvested in Buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.15% Nonidet P40(NP40) and 0.1 mM EGTA) supplemented with 1 mM DTT and PIC. Cells were homogenized through a 26G needle, nuclei were isolated by centrifugation and the supernatant (cytoplasmic fraction) was collected. Nuclei were washed quickly with ice-cold PBS, suspended in Buffer B (20 mM HEPES pH 7.9, 400 mM NaCl, 1 mM EDTA, 1 mM EGTA and 0,5% NP 40) supplemented with 1 mM DTT and PIC and lysed by sonication.

Protein concentration was determined using BCA Protein Assay Kit (Pierce). Equal amounts of protein were separated by SDS-PAGE and transferred onto nitrocellulose membrane (Amersham). After transfer, the membrane was blocked for one hour at room temperature in 5% skimmed milk in TBST buffer (25 mMTrisHCl, 137 mMNaCl, 0.1% Tween 20, pH 7.5) and probed with different antibodies. Incubation with primary antibody was performed overnight at 4°C followed by appropriate secondary HRP- conjugated antibodies (anti-rabbit

IgG or anti-mouse IgG, 1:5000, Amersham) for one hour at room temperature. Enhanced chemoluminescence (ECL Plus, Amersham) was used for detection of the bands. The following primary antibodies were used: mouse monoclonal anti-BMI1 (1:1000, clone AF27, Active Motif), anti-GAPDH (1:1000, G8795, Sigma), anti-mCherry (1:1000, ab125096, abcam), anti α -tubulin (1:5000, clone DM1A, Sigma) and anti-Vinculin (1:5000, V4505, Sigma); rabbit monoclonal anti-p44/42 MAPK (1:1000, 137F5, Cell Signaling) and anti-phospho-p44/42 MAPK (1:1000, D13.14.4E, Cell Signaling); rabbit polyclonal anti β -tubulin (1:4000, ab6046, abcam), anti-CHD7 (1:500, ab117522, abcam), anti-DUSP4 (1:500, ab72593, abcam) and anti-HA (1:1000, ab9110, abcam); goat polyclonal anti-Lamin B (1:5000, sc-6216, Santa Cruz).

Assay for Transposase-Accessible Chromatin (ATAC)

Intact nuclei were extracted from freshly isolated cells (50,000) of each condition (Ctr, shCHD7, shBMI1, shCHD7;BMI1) and ATAC libraries were produced as previously described (Buenrostro et al., 2013). Three independent biological replicates were used for each condition.

qPCR analyses were performed on a Stratagene Mx3000P thermal cycler (Agilent Technologies) using PrecisionPlus 2X Mastermix (Primerdesign) following the manufacturer's guidelines. The following primer pairs were used:

FW1: GAAGAGGCGGACCCAGCGGT, Rev1: TTCCTGCCGGTCATCTCGCTT (KAT6B);

FW2: GCGAGGAAGAGAAGAGAACCCG, Rev2: GCTCCGACTGCTATGTGACCG (433-259 down);

FW3: AAAGGAAATAGCCGGCTGAGGA, Rev3: AGTCTGGGGCTAGGAGGTGT (404-576 up);

FW4: AGGACTCGCTCGCAGTTTCG, Rev4: AAGGTCTCGGAAGTGGAGGCTC (1270 – 1509 up).

Technical triplicates for each sample were analysed. The Ct values of all experimental primer pairs were normalised against the Ct value of a control locus where DNA accessibility is known to be constitutively high (KAT6B), and then expressed as a percentage of the control group, with control being 100%.

Gene expression analysis of cell line BMI1 / CHD7 knockdowns

RNA sequencing was performed using the Illumina HiSeq2000 platform with poly(A) enrichment and paired end reads. Two biological replicates of the following conditions were used: Ctrl, shBMI1, shCHD7, shCHD7;shBMI1. Between 40 and 60 million reads were obtained in all cases. Illumina's BaseSpace was used to perform TopHat (Trapnell et al., 2009) alignments of all reads, and to run CuffLinks (Trapnell et al., 2010) for gene expression estimation followed by differential expression analysis. Log2 fold changes between conditions (Ctrl vs shBMI1, Ctrl vs shCHD7 and Ctrl vs shCHD7;shBMI1) were determined and genes differentially expressed in the replicates were determined as having a log2 fold change $> \pm 0.5$, and those with a p-value < 0.05 were further analysed.

Gene expression analysis of BMI1^{High};CHD7^{Low} patients

A z-score for BMI1 and CHD7 gene expression was produced for all patients in the cohort (Cho et al., 2011). Patients were classified as having a BMI1^{High};CHD7^{Low} signature if the BMI1 z-score was > 0.2 and if the CHD7 z-score was < -0.2 . The mean expression of every gene for patients in the BMI1^{High};CHD7^{Low} group vs the remaining group of patients was determined in order to evaluate differential expression. A permutation test was performed (n=1000) to assign a p-value for the observed difference in expression. Differentially expressed genes were cross-referenced with differentially expressed genes in the shCHD7 cells vs shCHD7;shBMI1 cells. Genes that agreed as being either upregulated or

downregulated in both cohorts, and were significant in both cohorts ($p < 0.05$) were shortlisted for further analysis.

Bioinformatic analyses

Pre-processing gene expression data from microarrays

Microarray data were subjected to the same pre-processing steps in all cases. Raw microarray data were first normalised using the RMA transformation (Irizarry et al., 2003) then the probe set annotation for the appropriate platform was used to assign probe sets to genes. Where multiple probe sets matched a single gene, the maximum value was used.

Pre-processing gene expression data from RNA sequencing

The three primary MB lines considered in this study (ICb1299, ICb1595 and CHLA-01-Med), in addition to the ICb1299 lentiviral knockdown conditions, were characterised by RNA sequencing. The raw data were manually checked with FastQC (Babraham Institute) to ensure that the reads were of sufficient quality and that no adapter sequences were found. The ICb1299 conditions were aligned to the Ensembl human reference genome GRCh38 using HISAT2 (Kim et al., 2015) followed by processing with featureCounts (Liao et al., 2014) to obtain gene counts. In the case of ICb1595 and CHLA-01-Med we used STAR (Dobin et al., 2013) to align to the same reference genome and obtain gene counts simultaneously.

Principal component analysis

Fig 3A shows the result of applying principal component analysis (PCA) to a published microarray dataset comprising 73 MB samples that have previously been classified by subgroup (Robinson et al., 2012) and the three cell lines described in this study (ICb1299, ICb1595, CHLA-01-Med).

To facilitate a consistent comparison across different gene expression measurement platforms, we first applied the YuGene transformation (Le Cao et al., 2014) to the Robinson microarray data. We then computed the PCA transformation, retaining the top 3 components. Subgroup centroids in the three-dimensional PCA coordinate system were calculated by assuming that the data follow a multivariate Gaussian mixture model and computing the ellipsoid enclosing 99% of the density.

We applied the same PCA transformation to YuGene-transformed gene count data obtained from RNA sequencing performed on the three MB cell lines included in this study. The ICB1595 sample is readily identifiable as belonging to G3. ICB1299 and CHLA-01-Med both fall within the boundary region between subgroups G3 and G4.

Unsupervised hierarchical clustering

As an additional verification of the molecular identity of the lines ICB1299 and ICB1595 we performed unsupervised hierarchical clustering of the tumour tissue and early passage gene expression data from these samples (Zhao et al., 2012) together with the Robinson cohort (Robinson et al., 2012) and the ICB1299, ICB1595 and CHLA-01-Med cell lines from this study (Fig. S3A). The Robinson and Zhao microarray data were prepared as previously described. Following the addition of a small offset value of 10^{-12} to the TPM values of the RNASeq-derived gene counts, the values were log2 transformed. Clustering was performed using the 100 genes identified by Northcott et al. as being highly differentially expressed between medulloblastoma subgroups (Northcott et al., 2012).

Gene expression heatmaps

Fig. S3C shows gene expression data from the knockdowns in the current study and nine healthy adult cerebellum samples, obtained from the Allen Human Brain Atlas (AHBA)

(Hawrylycz et al., 2012). Following the addition of a small offset value of 10^{-12} to the TPM values, both datasets were log2 transformed. The data were standardised by subtracting the AHBA sample mean and dividing by the AHBA sample standard deviation for each gene. We display gene expression values corresponding to 50 genes previously implicated in medulloblastoma G3 and G4 membership (Northcott et al., 2012).

References

- Brett, B. T., Berquam-Vrieze, K. E., Nannapaneni, K., Huang, J., Scheetz, T. E., and Dupuy, A. J. (2011). Novel molecular and computational methods improve the accuracy of insertion site analysis in Sleeping Beauty-induced tumors. *PloS one* 6, e24668.
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., and Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods* 10, 1213-1218.
- Cho, Y. J., Tsherniak, A., Tamayo, P., Santagata, S., Ligon, A., Greulich, H., Berhoukim, R., Amani, V., Goumnerova, L., Eberhart, C. G., *et al.* (2011). Integrative genomic analysis of medulloblastoma identifies a molecular subgroup that drives poor clinical outcome. *J Clin Oncol* 29, 1424-1430.
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* 29, 15-21.
- Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., van de Lagemaat, L. N., Smith, K. A., Ebbert, A., Riley, Z. L., *et al.* (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* 489, 391-399.
- Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249-264.

Khodosevich, K., Inta, D., Seeburg, P. H., and Monyer, H. (2007). Gene expression analysis of in vivo fluorescent cells. *PloS one* 2, e1151.

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature methods* 12, 357-360.

Le Cao, K. A., Rohart, F., McHugh, L., Korn, O., and Wells, C. A. (2014). YuGene: a simple approach to scale gene expression data derived from different platforms for integrated analyses. *Genomics* 103, 239-251.

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)* 30, 923-930.

Merve, A., Dubuc, A. M., Zhang, X., Remke, M., Baxter, P. A., Li, X. N., Taylor, M. D., and Marino, S. (2014). Polycomb group gene BMI1 controls invasion of medulloblastoma cells and inhibits BMP-regulated cell adhesion. *Acta neuropathologica communications* 2, 10.

Morrissy, A. S., Garzia, L., Shih, D. J., Zuyderduyn, S., Huang, X., Skowron, P., Remke, M., Cavalli, F. M., Ramaswamy, V., Lindsay, P. E., *et al.* (2016). Divergent clonal selection dominates medulloblastoma at recurrence. *Nature* 529, 351-357.

Northcott, P. A., Shih, D. J., Peacock, J., Garzia, L., Morrissy, A. S., Zichner, T., Stutz, A. M., Korshunov, A., Reimand, J., Schumacher, S. E., *et al.* (2012). Subgroup-specific structural variation across 1,000 medulloblastoma genomes. *Nature* 488, 49-56.

Robinson, G., Parker, M., Kranenburg, T. A., Lu, C., Chen, X., Ding, L., Phoenix, T. N., Hedlund, E., Wei, L., Zhu, X., *et al.* (2012). Novel mutations target distinct subgroups of medulloblastoma. *Nature* 488, 43-48.

Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics (Oxford, England)* 25, 1105-1111.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals

unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* 28, 511-515.

Zhao, X., Liu, Z., Yu, L., Zhang, Y., Baxter, P., Voicu, H., Gurusiddappa, S., Luan, J., Su, J. M., Leung, H. C., and Li, X. N. (2012). Global gene expression profiling confirms the molecular fidelity of primary tumor-based orthotopic xenograft mouse models of medulloblastoma. *Neuro-oncology* 14, 574-583.